

Why

Understanding the probability rules is important for both understanding the language necessary for stating statistical results and understanding the way samples are related to populations - the basis of statistical inference.

Conditional probability deals with the situation in which the possible outcomes of a random process are somehow restricted (we have found out part—but not all— of the result, or we sample from only part of the population, conditions have changed, etc.) and we want the probabilities for this modified process. For experimental/theoretical work, this is particularly important for deciding whether two events are *independent* or are *related* (dependent). The basic multiplication rule for probabilities (used with “and” situations) is based on conditional probability, with a special case for independent events. *Tree diagrams* are an analytical tool that is useful for keeping track of *sequential* ideas and conditional probabilities.

LEARNING OBJECTIVES

1. Learn to use the basic probability rules in combinations in different settings.
2. Be able to use a tree diagram for a multi-step probability question.
3. Continue developing skills in working with teams.

CRITERIA

1. Success in working as a team and in fulfilling the team roles.
2. Success in involving all members of the team in the conversation.
3. Success in completing the exercises.

RESOURCES

1. Your text, sections 5.3 and 5.4
2. The team role desk markers (handed out in class for use during the semester)
3. Your class notes from Monday and Wednesday
4. Your calculator
5. 40 minutes

PLAN

1. Select roles, if you have not already done so, and decide how you will carry out steps 2 and 3 (5 minutes)
2. Work through the exercises given here - be sure everyone understands all results & procedures(30 minutes)
3. Assess the team’s work and roles performances and prepare the Reflector’s and Recorder’s reports including team grade (5 minutes).

FORMULAS

Standard probability formula for conditional probability: $P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$

General multiplication rule for probabilities: $P(E \text{ and } F) = P(E) * P(F|E) = P(F) * P(E|F)$

Addition rule for probabilities: $P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$

Complement rule for probabilities: $P(E^c) = 1 - P(E)$

A special addition rule [“partition rule”]: $P(E) = P(E \text{ and } F) + P(E \text{ and } F^c)$

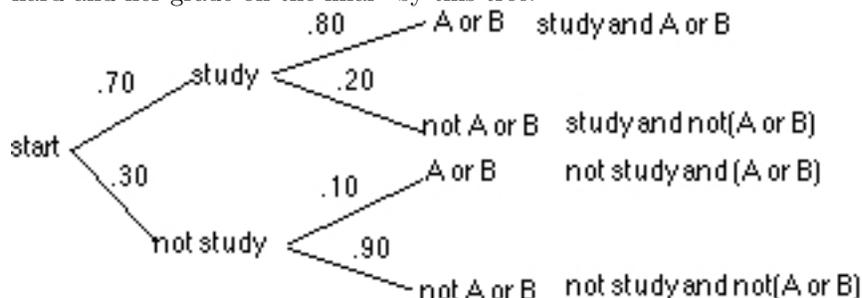
Complements and conditional probabilities: $P(E^c|F) = 1 - P(E|F)$

Independence (and relationship): Two events E and F are *independent* if $P(E|F) = P(E)$ [conditional probability for E = unconditional probability for E]. In this case, $P(E \text{ and } F) = P(E) * P(F)$. Otherwise they are *related* (but we may not know why).

TREE DIAGRAMS

A *tree diagram* is useful for keeping track of cases and steps when both are involved in a program or when conditional probability and cases are both involved. The tree *branches* whenever there are two possible next steps in a situation, and the branches are marked with the *conditional* probabilities. Each outcome is represented by a path from start to finish through the probability of the outcome is obtained by multiplying the probabilities shown on the branches along the path.

An example: Suppose 30% of the students in an advanced snoring class study hard (and the others do not), and suppose that 80% of those who study hard get A or B on the final exam, while only 10% of the others get A or B. We can represent the outcomes for “We choose a student at random and observe whether she studies hard and her grade on the final” by this tree:



Then the result “We get a student who studies and gets an A or B” is represented by the first path through the tree, and its probability is $P(\text{study and (A or B)}) = P(\text{studies}) * P(\text{A or B}|\text{studies}) = .7 * .8 = .56$, while “We get a student who does not study but still gets A or B” is represented by the third path and its probability is $P(\text{not study and (A or B)}) = .3 * .1 = .03$. The probability of getting a student who gets an A or B is obtained by the “partition rule” given above: $P(\text{A or B}) = P(\text{study and (A or B)}) + P(\text{not study and (A or B)}) = .56 + .03 = .59$. (59% of students get A or B—56% are successful studiers, 3% are lucky non-studiers)

EXERCISE

- At the East Overshoe State, 70% of the students receive financial aid, 60% of the students are commuters, and 80% of the commuters receive financial aid [Notice that this 80% is a *conditional* probability because it is “of the commuters”].
 - If we pick a student at random, what is the probability we will get a student who is a commuter and receives financial aid [This is *joint* probability—not conditional—it *asks for* both conditions] ?
 - If we pick a student at random, what is the probability we will get a student who is a commuter or receives financial aid?
 - If we pick a student at random, what is the probability we will get a student who is a commuter and does not receive financial aid?
 - If we pick a student at random, are the events “commuter” and “financial aid” independent? Explain (using the probabilities)
- Here is a table with information on the 40 students enrolled in another course at the College of Knowledge

		Major				Totals
		English	Theater	Philosophy	Other	
Year	Sophomore	6	4	5	2	17
	Junior	5	4	0	7	16
	Senior	1	2	2	2	7
Totals		12	10	7	11	40

- If we select one of these students at random, what is the probability of getting a Junior?
 - If we select a Philosophy major at random, what is the probability of getting a Junior?
 - Are the events “Get a Junior” and “Get a Philosophy Major” independent? Explain.
 - Are the events “Get a Junior” and “Get a Theater major” independent? Explain.
 - If *two* students (two distinct students) are chosen, what is the probability of getting exactly one Senior [either the first or the second is a senior - but not both]?
- Why is it impossible for two disjoint events (with probability that is not 0) to be independent?

4. (You will want to draw a tree diagram for this problem) Screening tests for medical conditions (AIDS, exposure to tuberculosis, sickle cell anemia, etc.) are not perfect. Their effectiveness is described by two conditional probabilities: The *sensitivity* is $P(\text{test is positive} \mid \text{person has condition})$ [probability the test detects the condition if applied to a person with the condition].

The *specificity* is $P(\text{test is negative} \mid \text{person does not have the condition})$ [probability the test detects “health” if applied to a person who does not have the condition].

These are the *conditional* probabilities for being right — depending on which answer is right. The “absolute” (unconditional) probability of being right depends on the proportion of people with the condition.

Suppose that 15% of the people in a state are susceptible to heebie-jeebies and our team of medical experts has developed a test for susceptibility to heebie-jeebies with 90% sensitivity (will be positive for 90% of cases) and 80% specificity (will be negative for 80% of non-cases). Represent this situation with a tree diagram—first step is “person is susceptible Y/N” second step is “test says person susceptible Y/N” so there are four branches by the end.

- (a) If we administer the test to a randomly chosen person, what is the probability that we have picked a person who is susceptible and the test will detect it (will be positive)?
- (b) If we administer the test to a randomly chosen person, what is the probability that we have picked a person who is not susceptible but (and) the test says the person is susceptible (test is positive even though person is not susceptible)
- (c) If we administer the test to a randomly selected person, what is the probability that the test will say the person is susceptible? (You need the results from the “partition rule” given above. The people with positive test results are divided into two groups—“susceptible” and “not susceptible”)
- (d) Suppose we administer the test to a randomly selected person and the test says the person is susceptible. What is the probability the person really is susceptible (That is, $P(\text{person is susceptible} \mid \text{test is positive})$)? Most people find this result surprising—Even with high sensitivity and specificity, if the condition is relatively rare, positive test results indicate *increased likelihood* of the condition but the probability may still be low. This is a reason for doing multiple independent tests.

READING ASSIGNMENT (in preparation for next class)

Section 6.1 on discrete random variables

SKILL EXERCISES:(hand in - individually - at next class meeting)

Sullivan p.254 #18-19, p.286 (Review Exercises) #6-10, 16, 17 - 19 (for 17 assume independence for results of complaints, use complement rules)