Activity: Basic Probability Rules

Concepts: Basic Probability Rules (does not include conditional probability)

Prerequisites: Students should be familiar with basic set theory, the probability function, sample spaces and events, and computing probabilities using the "equal likeliness" principle.

Goals:
- To lead students to discover and understand some basic rules of probability:
  - Complement rule
  - Addition rule for union of two events
  - Addition rule for disjoint events
  - Addition rule for finite union of events
- To introduce students to common notation for events
- To provide experience with translating probability questions into event/set terminology and notation, and then solving the problems using the above rules.

Recap: Recall the set notation A', $A \cup B$, and $A \cap B$ on the **Brief Review of Set Operations and Properties** handout and in Appendix A.1 of your text. Always keep in mind that events are sets (in particular, they are subsets of the sample space S). Thus, it is legitimate to perform set operations such as complement, intersection, and union on them. On the other hand, probabilities are numbers. Thus, it is legitimate to add, multiply, and divide probabilities but not to take complements, intersections, or unions of them.

Situation: In 1998 the American Film Institute created a list of the top 100 American films ever made (www.afi.com/tv/movies.asp). Suppose that three people gather to watch one of these movies and, to avoid potentially endless debates about a selection, decide to choose a movie at random from the "top 100" list. You will investigate the probability that it has already been seen by at least one of the three people.

Let A denote the subset of these 100 films that Allan has seen, so the event A = {films that Allan has seen}. Similarly define events B and F for Beth and Frank, respectively. The following 2x2 table classifies each movie according to whether it was seen by Allan and whether it was seen by Beth. The "at random" selection implies that each of the 100 films is equally likely to be chosen; i.e., each has probability 1/100. Thus, the probabilities of these various events can be calculated by counting how many of the 100 films comprise the event of interest.

	Beth yes	Beth no	Total
Allan yes	42	6	
Allan no	17	35	
Total			100

For example, the table reveals that 42 movies were seen by both Allan and Beth, so $P(A \cap B)=42/100=.42$ (read "the probability of A and B").

1. Translate the following events into set notation using the symbols A and B, complement, union, intersection. Also give the probability of the event as determined from the table:

Event in words	Event in set notation	Probability
Allan and Beth have both seen the film	$A \cap B$	$P(A \cap B)=.42$
Allan has seen the film and Beth has not		
Beth has seen the film and Allan has not		
Neither Allan nor Beth has seen the film		

2. Fill in the marginal totals of the table (the row and column totals). From these totals determine the probability that Allan has seen a randomly selected film and also the probability that Beth has seen the film. (Remember that the film is chosen at random, so all 100 are equally likely.) Record these, along with the appropriate symbols, below.

   P(Allan has seen it) = P(   ) =
   P(Beth has seen it) = P(   ) =

3. Determine the probability that Allan has not seen the film. Do the same for Beth. Record these, along with the appropriate symbols, below.

4. If you had not been given the table, but instead had merely been told that $P(A)=.48$ and $P(B)=.59$, would you have been able to calculate $P(A')$ and $P(B')$? Explain how.

One of the most basic probability rules is the *complement rule*, which asserts that the probability of the complement of an event equals one minus the probability of the event: $P(A') = 1-P(A)$

5. *Add the counts in the appropriate cells of the table* to calculate the probability that at least one of Allan or Beth has seen the film (either Allan or Beth or both). Also indicate the symbols used to represent this event.

6. If you had not been given the table but instead had merely been told $P(A)=.48$ and $P(B)=.59$ and were asked to calculate $P(A \cup B)$. You might first consider $P(A)+P(B)$. Calculate this sum and compare the result to your answer to 5). Are they the same? Is this even a legitimate answer?

7. Is P(A)+P(B) larger or smaller than P(A $\cup$ B)? By how much? Explain why this makes sense, and indicate how to adjust P(A)+P(B) to determine P(A $\cup$ B).

The *addition rule* asserts that the probability of the union of two events can be calculated by adding the individual event probabilities and then subtracting the probability of their intersection: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

8. Use this addition rule as a second way to calculate the probability that Allan or Beth has seen the movie, verifying your answer to 5).

9. As a third way to calculate P(A $\cup$ B), first identify (in words and in symbols) the complement of the event {Allan or Beth has seen the movie}. Then find the probability of this complement *from the table*. Then use the complement rule to determine P(A $\cup$ B). Are your answers to 5) and 8) confirmed?

10. What would have to be true about A and B for it to be valid to say that $P(A \cup B) = P(A) + P(B)$? *Hint*: What would the Venn Diagram look like in this case?

Two events A and B are said to be *disjoint* (or *mutually exclusive*) if their intersection is the empty set ϕ. In other words, two events are disjoint if they cannot both happen simultaneously. If $A \cap B=\phi$, then it follows that $P(A \cup B) = P(A) + P(B)$. This is known as the *addition rule for disjoint events*; it is a special case of the addition rule since if $A \cap B = \phi$, $P(A \cap B) = P(\phi) = 0$.

<u>Three Events:</u>

Now consider Frank, who has seen 61 of the top 100 films. The following pair of 2x2 tables (which can be considered a 2x2x2 table) reveal the counts. The 2x2 table on the left pertains to films Allan has seen, and the 2x2 table on the right pertains to films Allan has not seen:

Allan yes				Allan no			
	Frank yes	Frank no	Total		Frank yes	Frank no	Total
Beth yes	39	3	42	Beth yes	14	3	17
Beth no	3	3	6	Beth no	6	29	35
Total	42	6	48	Total	20	32	52

11. What is the probability that all three of them have seen the film? Also indicate the symbols used to denote this event.

12. Express in symbols and in words the event for which the probability can be read directly from the table to be 14/100.

We will again find three ways to calculate the probability that at least one of these three people has seen the film.

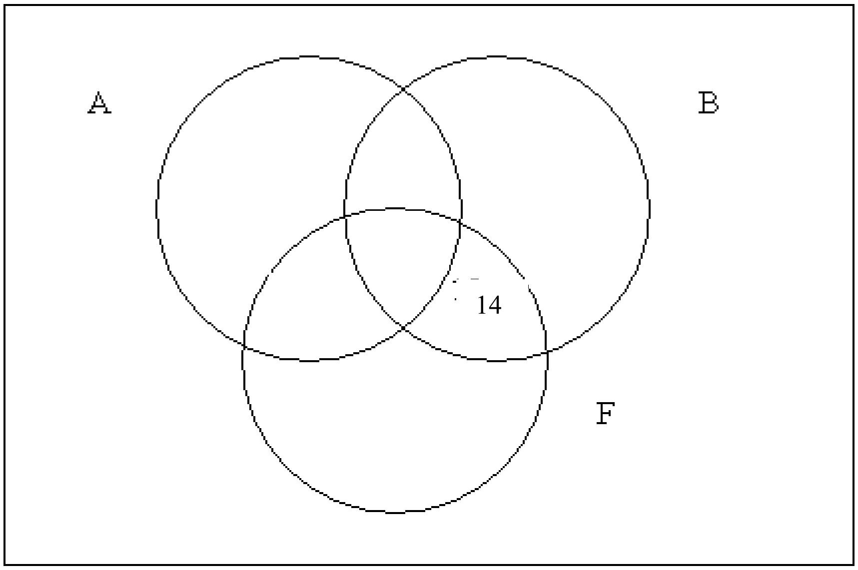
13. Express this event (that at least one of these three people has seen the film) in symbols.

14. Determine $P(A \cup B \cup F)$ directly from the table by adding the counts of the outcomes that comprise this event.

15. Express the complement of this event (that at least one of these three people has seen the film) in words and in symbols. Use the table to determine the probability of this complement. Then use the complement rule to find the probability that at least one of the three has seen the film. Does the answer agree with 14)?

16. Suppose that instead of being given the tables, you had only been told that Allan has seen 48 of the films, Beth 59, and Frank 62. Would that information alone enable you to determine the number of films that at least one of these three has seen? Explain.

To see how to calculate $P(A \cup B \cup F)$ from other probabilities, consider the Venn diagram below. The upper left circle represents movies that Allan has seen, the upper right circle for Beth, and the bottom circle for Frank (note that the circles and overlaps are not drawn to scale).



17. Take the eight numbers (counts) from the cells of the 2x2x2 table and insert them into the appropriate regions of the Venn diagram. For example, to get you started, the "14" representing films seen by Beth and Frank but not Allan has already been inserted.

18. Use the Venn diagram to figure out how to calculate $P(A \cup B \cup F)$ from $P(A)$, $P(B)$, $P(F)$, $P(A \cap B)$, $P(A \cap F)$, $P(B \cap F)$, and $P(A \cap B \cap F)$. Write out the rule that you propose, and verify that it gives the correct answer in this case. [Note, for instance, that the event $A \cap B$ includes both outcomes $A \cap B \cap F$ and $A \cap B \cap F'$.]

You have discovered the *addition rule for three events*:
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$