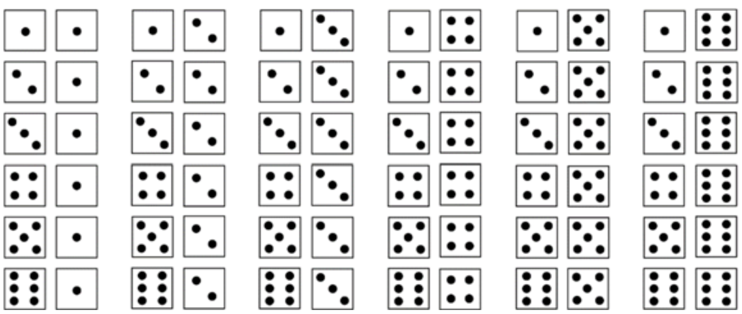
If you roll a die, what is the probability you get a four OR a five? What is the probability that you get neither?

General Education Statistics
Class Notes
The Addition Rule (for "OR" events) and Complements (Section 5.2)

Consider rolling two dice. How likely is it that you get two even numbers OR a sum of 5? What about two even numbers OR a sum of 6? How would you find these probabilities? Let's investigate these and other probabilities.

What does OR mean?
This may seem obvious but we use OR in math a little differently than in English. When we talk about event E OR event F, we mean either E or F occurs, or *possibly both*, occur.

expl 1: Consider rolling two distinguishable, fair, six-sided dice. The sample space is given below. Answer the following questions. This will help us understand an important formula we will see later.



Recall the probability of an event is the number of successes divided by the number of possibilities.

a.) What is the probability of getting a sum of 5? Circle the successes above.

b.) What is the probability of getting a sum of 7? Circle the successes above.

c.) What is the probability of getting a sum of 5 OR a sum of 7? Notice how you have already circled all of the successes.

d.) Can the events "sum of 5" and "sum of 7" happen at the same time (in one roll of the dice)?

Definition: Two events are **disjoint** if they have *no* outcomes in common. This means that they *cannot occur at the same time*. These are often called **mutually exclusive** events.

This is true of the events "sum of 5" and "sum of 7". We have a very important rule that applies to mutually exclusive events, which we saw at work in example 1.

Addition Rule for Disjoint Events:
If E and F are disjoint (or mutually exclusive) events, then $P(E \text{ or } F) = P(E) + P(F)$.

Again, why are "sum of 5" and "sum of 7" considered disjoint?

You may see this written as $P(E \cup F)$. This is the **union** symbol.

expl 2: Let's take this idea further. Consider the events "sum of 5", "sum of 7", "sum of 6", "sum of 4", and "sum of 8". Do these events share any outcomes in common? In other words, are they mutually exclusive? What, do you think, can we say about how we should find P(sum of 5 OR sum of 7 OR sum of 6 OR sum of 4 OR sum of 8)?

We can see how the above rule may be generalized for any number of mutually exclusive events.

Addition Rule for Many Disjoint Events:
If $E, F, G, H, \ldots$ are disjoint (or mutually exclusive) events, then

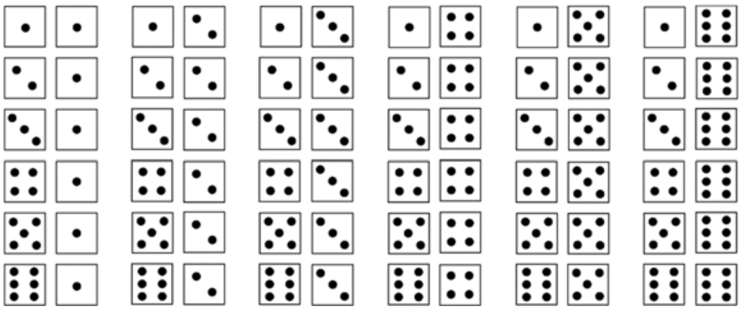
$$P(E \text{ or } F \text{ or } G \text{ or } H \text{ or } \ldots) = P(E) + P(F) + P(G) + P(H) + \ldots .$$

To all be disjoint, no two events could overlap.

Do you see this at work in example 2?

But how do we find probabilities of events that are *not* mutually exclusive? Let's consider this new example.

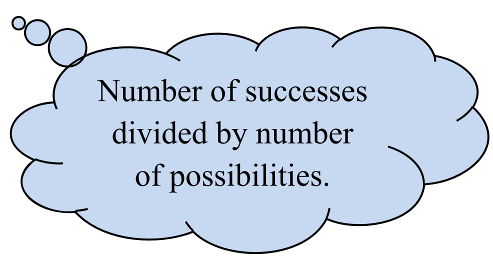
expl 3: Again, consider the two dice as before. Answer the following questions.



a.) What is the probability of both dice coming up even, that is P(both even)? Circle the successes above. Do *not* reduce.

b.) What is the probability of getting a sum of 6, that is P(sum is 6)? Circle the successes above. Do *not* reduce.

c.) What is the probability of both dice coming up even OR a sum of 6? Recall, this means that one, or the other, or possibly both events occur. Return to the sample space and actually count the successes to use the Very Useful Formula for probability. Do *not* reduce.



expl 3 continued:

d.) The answer to part *c* is 12/36. If we simply add *P*(sum is 6) + *P*(both even), we do *not* get 12/36. What do we get and why is this *not* equal to the probability we are after in part *c*? In other words, what are we counting that we should *not* be?

e.) To further clarify, the outcomes that were overcounted in part *d* satisfy “sum is 6” AND “both even”? In other words, those outcomes satisfy *both* of the events.
Calculate *P*(sum is 6 AND both even). Do *not* reduce.

AND means *both* events occur.

When we add the probabilities as in part *d*, we double-count the outcomes in this intersection.

This overcount needs to be corrected by subtracting those outcomes here.

f.) Use the probabilities you have found to verify this equation.
P(sum is 6 OR both even) = *P*(sum is 6) + *P*(both even) – *P*(sum is 6 AND both even)

This leads us to a very important rule.

> You may see this written as $P(E \cap F)$. This is the **intersection** symbol.

The General Addition Rule:
For any two events E and F, we know $P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$.

Since counting the possibilities and successes is a large part of calculating probabilities, you will also see these formulas in terms of the *number of outcomes* in events and *not* their probabilities.

This can work for any formula in this section, but I show it here only for the General Addition Rule. For instance, you may see $N(E \text{ or } F) = N(E) + N(F) - N(E \text{ and } F)$, where $N(E)$ means the number of outcomes in event E, etc.

> Same idea, but uses number of outcomes instead of probabilities.

Optional Worksheet: Probability: Addition Rule:
This worksheet gives us a quick example of the General Addition Rule. We will use experimental (or empirical) probability to explore the rule.

Worksheet: Counting problems involving "OR":
This worksheet explores counting the successes for probability questions involving OR. We will look at when the events do have outcomes in common and when they do *not*. Since we can often use the fundamental definition for probability (number of successes divided by number of possibilities), being able to count the successes is vitally important.

Contingency Tables:

Definition: A **contingency table** or **two-way table** shows two categories of data and how they interact. Surveys, whose respondents are broken into categories, often result in two-way tables. There will be a **row variable** and a **column variable**. Each number is said to be in a **cell** in the table.

expl 4: Below is a contingency table that shows the relationship between cigar smoking and dying from cancer. Answer the questions that follow.

	Died from Cancer	**Did Not Die from Cancer**	**Total**
Never smoked cigars	782	120,747	
Former cigar smoker	91	7,757	
Current cigar smoker	141	7,725	
Total			

(source: National study of 137,243 U.S. men, *Journal of the National Cancer Institute*, Feb. 16, 2000)

a.) Find and record the totals for each column and each row. This is the first thing you should do with such a table.

b.) If a survey respondent was chosen at random, what is the probability that he was a former cigar smoker?

c.) If a survey respondent was chosen at random, what is the probability that he died from cancer?

d.) If a survey respondent was chosen at random, what is the probability that he was a former cigar smoker AND died from cancer?

What is considered a success? How many successes are there? How many possibilities?

expl 4 continued:
e.) If a survey respondent was chosen at random, what is the probability that he was a former cigar smoker OR died from cancer?

f.) If a survey respondent was chosen at random, what is the probability that he was *not* a former cigar smoker?

Definition: Complement: Let S denote the sample space of a probability experiment and let E denote an event. The complement of E, denoted E^C, is all outcomes in the sample space S that are *not* outcomes in the event E.

The events "former cigar smoker" and "*not* a former cigar smoker" are complements.

Consider rolling one six-sided die. Can you think of two events that are complements?

There are other notations used for the complement, like $\overline{E}$, $\sim E$, or E'.

expl 5: I have a bag with ten marbles: four red, three yellow, and three blue.
a.) If I randomly select one marble from the bag, what is the probability that it is blue?

b.) If I randomly select one marble from the bag, what is the probability that it is *not* blue?

How many are *not* blue?

expl 6: I will select a card from *only the face cards* from a poker deck. (The face cards are Jack of Hearts, Queen of Hearts, King of Hearts, Jack of Diamonds, Queen of Diamonds, King of Diamonds, Jack of Spades, Queen of Spades, King of Spades, Jack of Clubs, Queen of Clubs, and King of Clubs. Clubs and Spades are black, Hearts and Diamonds are red.)
a.) If I select one card from my *partial* deck, what is the probability that it is a black Queen?

b.) If I select one card from my *partial* deck, what is the probability that it is *not* a black Queen?

How many cards are *not* black Queens?

These probabilities involving complements can be done the traditional method of counting the successes and possibilities. However, we could also use a basic rule.

Complement Rule:
If E represents any event and E^C represents the complement of E, then $P(E^C) = 1 - P(E)$.

Since E and E^C make up all possibilities, their probabilities must add to 1.

Redo example 5*b* but use this formula. (This assumes you did *not* use the formula when you did the problem originally.)

expl 5*b* again: If I select one marble from the bag (with ten marbles: four red, three yellow, and three blue), what is the probability that it is *not* blue?