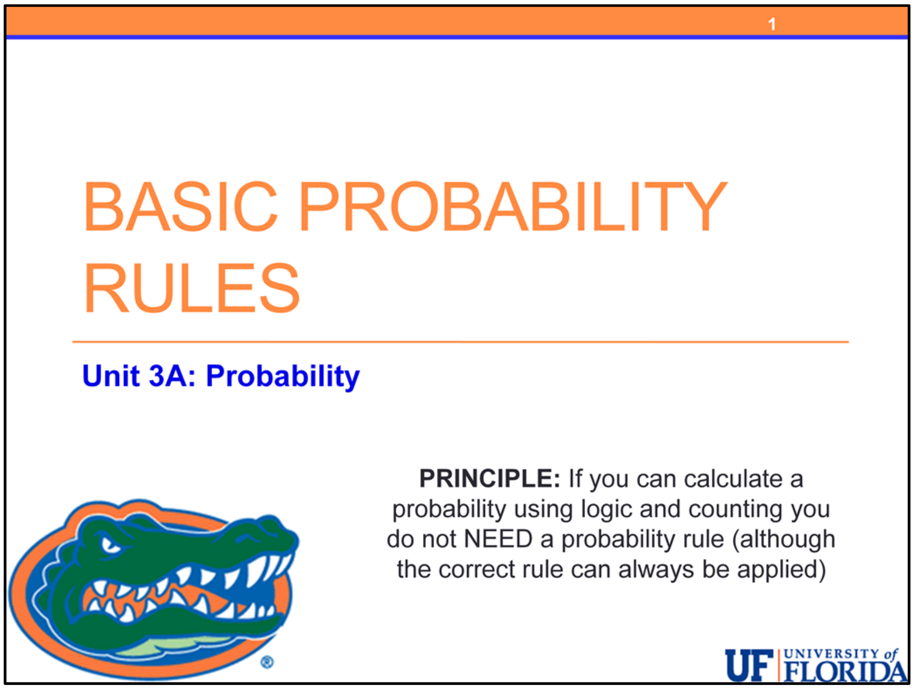
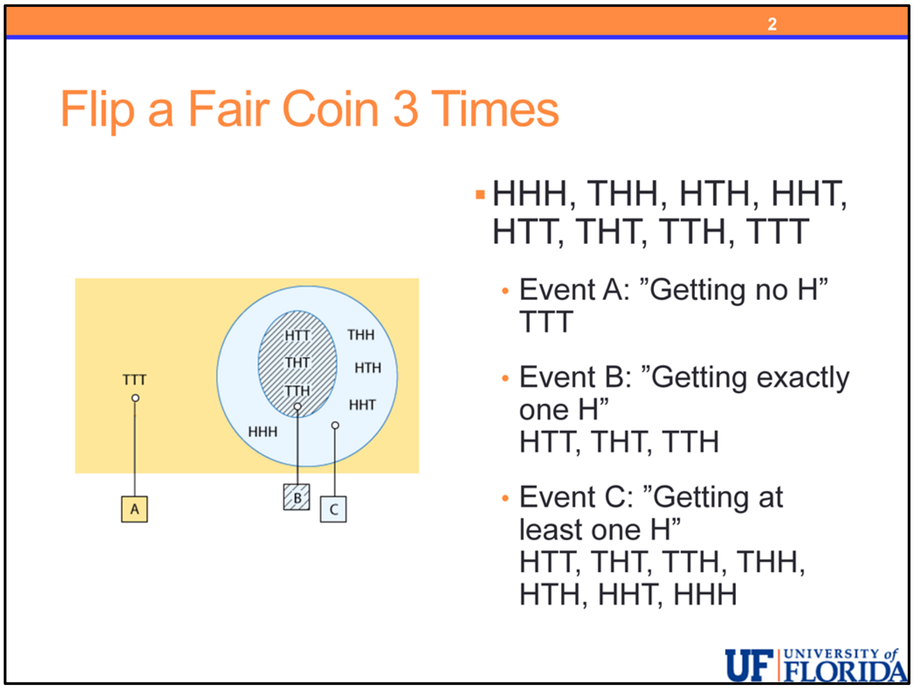


Now we will work a few problems, discussing some of the basic probability rules as we go.

Before we begin, I want to point out a theme we will repeat. If you can calculate a probability using logic and counting you do not NEED a probability rule (although the correct rule can always be applied).

We will illustrate a number of examples using both the logical method and that which uses the rule.



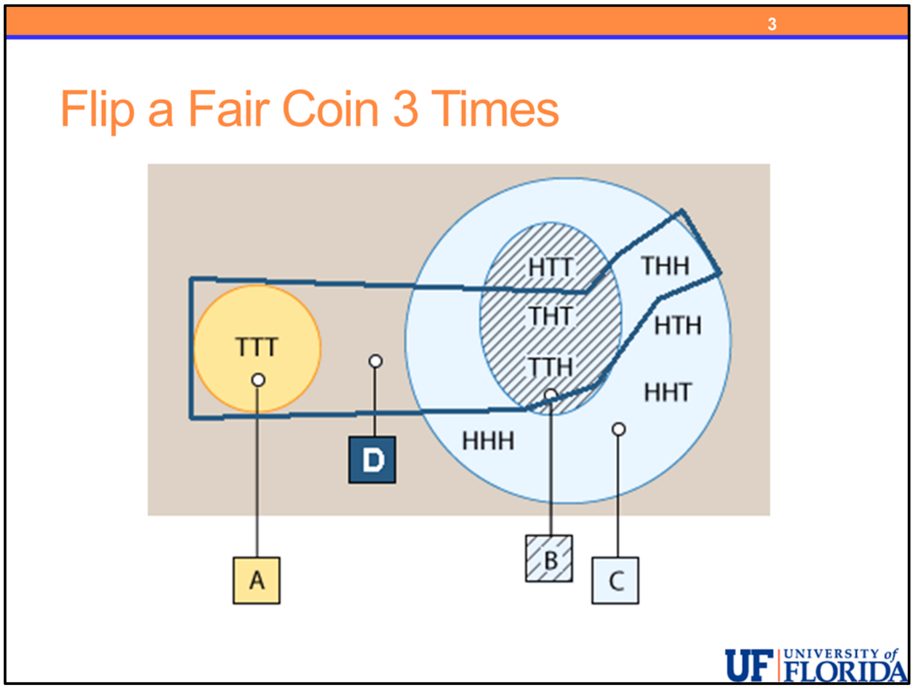
If we flip a fair coin 3 times, we have 8 possible outcomes: {HHH, THH, HTH, HHT, HTT, THT, TTH, TTT}.

By definition, an event corresponds to some collection (subset) of the possible outcomes. Let’s define the following events:
Event A: ”Getting no H”
Event B: ”Getting exactly one H”
Event C: ”Getting at least one H”

- We can see that event A “getting no heads” corresponds to TTT
- Event B “getting exactly one head” corresponds to one of three outcomes: HTT, THT, TTH
- And that Event C “getting at least one head” contains every outcome except TTT since all others have at least one head [explicitly they are HTT, THT, TTH, THH, HTH, HHT, HHH].

The Venn diagram shows these events with outlines, shading, and labels for each event.

It just happens that every outcome is counted in one of these events but this is not necessary in such a list.



If we add Event D: ”Getting a T on the first toss” → THH, THT, TTH, TTT

We can visualize it outlined in dark blue above.

Notice in this version there is a gray background that doesn’t reflect any events – only “empty space”

You can work through the activities in the notes to calculate some probabilities using this example. The solutions are given in the activities so we won’t go through them here.

But remember, that the only reason these 8 outcomes are equally likely is that we are talking about flipping a FAIR coin.

If the coin was unfair, these 8 outcomes would NOT be equally likely and we will learn more about how to handle that situation when we learn about binomial random variables in Unit 3B.

Example: Blood Types

- Probability Rule One:
  For any event A, 0 ≤ P(A) ≤ 1

Blood type	O	A	B	AB
Probability	0.44	?	0.10	0.04

We have already mentioned our first rule: Probability Rule One states that For any event A, the probability of the event is always between 0 and 1, i.e. 0 ≤ P(A) ≤ 1. We do include the possibility that P(A) = 0 or 1 although usually these events are not very interesting to study and discuss.

Here we have a table representing the probabilities of each blood type in the US Population.

How can we find the missing value?

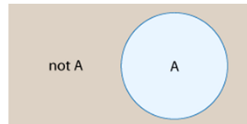
Logic tells us that, if these are all of the choices, then the missing value must be everything not accounted for by the other blood types.

We can find P(Blood Type A) by adding the other blood types to get 0.44 + 0.10 + 0.04 = 0.58 and then determining that P(A) must be 0.42 in order to account for everyone.

5

Example: Blood Types

- Probability Rule Two:
  Sum of the probabilities of all possible outcomes is 1
- Probability Rule Three
  (The Complement Rule):
  P(not A) = 1 – P(A)



Blood type	O	A	B	AB
Probability	0.44	0.42	0.10	0.04

UF UNIVERSITY of FLORIDA

The logic we just used is the basis for the next two rules.

Probability Rule Two: The sum of the probabilities of all possible outcomes is 1.

And Probability Rule Three (The Complement Rule): P(not A) = 1 – P(A);
that is, the probability that an event does not occur is 1 minus the probability that it does occur.

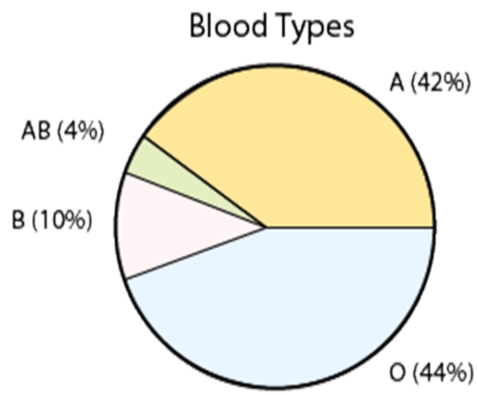
- This rule can also be written as P(A) = 1 – P(not A) or that P(A) + P(not A) = 1.

The complement rule can be useful when finding the probability of the "opposite" event is easier than the current event of interest.

For example: P(At least one head in 3 flips of a fair coin)
= 1 – P(Do not get at least one head in 3 flips of a fair coin)
= 1 - P(No heads in 3 flips of a fair coin).

It is very easy to calculate P(No heads) but it is more difficult to calculate P(At least one head) and so the complement rule can be helpful.

Example: Blood Types



We used this example in Unit 1. The only difference is that we were talking about percentages within a sample in Unit 1. Now we are USING those values from our sample as ESTIMATES of the “TRUE” probability.

Or sometimes, in our examples in this section on probability, we might present information as if we know the truth in the population of interest, which, to a certain extent we are doing now.

This pie chart represented the population in Unit 1 and now, the only change is that for probabilities we will use the decimal representation as opposed to a percentage.

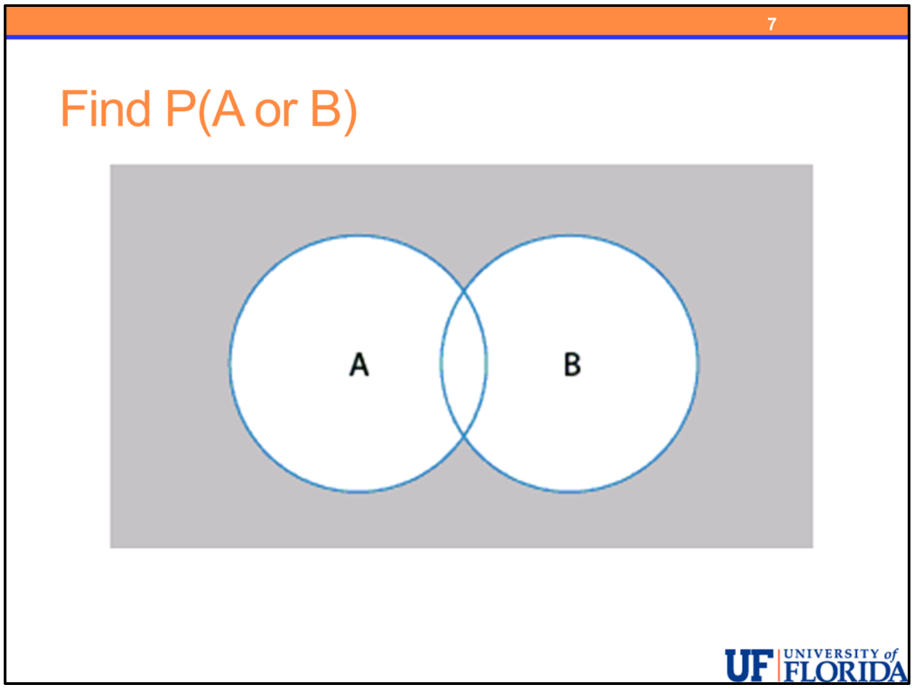
Before moving on, let’s talk a little about rounding.

In general you should carry probabilities to at least 4 decimal places for intermediate steps. We often round our final answer to two or three decimal places.

For extremely small probabilities, it is important to have 1 or two significant digits (non-zero digits after the decimal) [such as 0.000001 or 0.000034].

In addition, many computer packages might display extremely small values using scientific notation.

For this entire course, it is a good habit to keep at least 2 more decimal places in intermediate steps than you need for your final answer. This will almost always guarantee that we will all agree in our final answers.

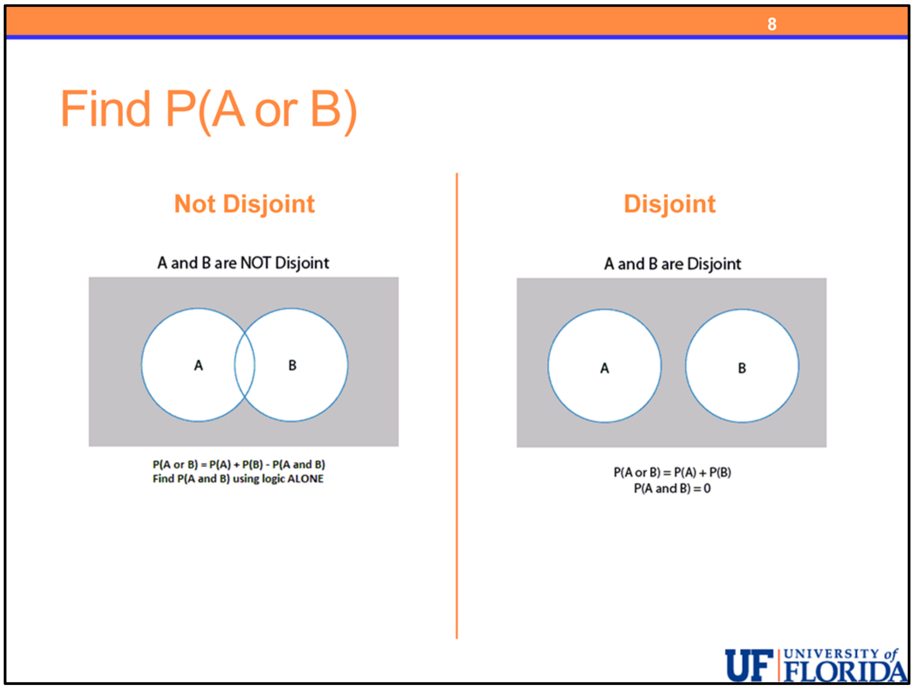


Our next two rules are for finding the probability of the compound event "A or B" – we must be careful as this is an inclusive "OR."

P(A or B) = the probability that A happens or B happens or both A and B happen.

I don't suggest using formulas to find P(A or B) unless it is necessary or if it really is easier for you. It is usually easy enough to determine this probability by logic and counting.

Before you can understand the difference between these two rules, you need to know what we mean by disjoint events.



Two events are disjoint if they cannot happen together – the events do not overlap – they have no intersection.

In the images above you can see that on the left, the two events A and B overlap, there is some probability that both A and B will occur. These two events are not disjoint.

On the right, there is no overlap. No possible outcomes satisfy both A and B simultaneously. The events A and B are disjoint.

The two rules we have for P(A or B) are:

Probability Rule Four (The Addition Rule for Disjoint Events) which says
If A and B are disjoint events, then P(A or B) = P(A) + P(B).

And Probability Rule Five (The General Addition Rule) which says:
For ANY pair of events A and B, P(A or B) = P(A) + P(B) – P(A and B).

Probability rule 4 is a specialized version of probability rule 5 since when A and B are disjoint, there is no intersection and the P(A and B) = 0 and so rule (#5) simplifies to P(A or B) = P(A) + P(B).

If you are not absolutely certain that two events are disjoint then it is best to use Rule 5, the General Addition Rule, as it will work for any pair of events. If you use this rule to calculate P(A or B), it will be best to calculate P(A and B) by logic alone. Notice as we complete our examples that we do NOT use any rule to calculate P(A and B), only logic.

Example – Disability vs. Age Groups

- Adults with health problem(s) that requires the use of special equipment

Age:	Yes	No
35-44	30	762
45-54	117	992
55-64	220	1428
65+	611	2422

The following example using data from a CDC survey of Florida residents age 35 and older will be used to illustrate the calculation of some basic probabilities.

Individuals were asked if they had any health problems that require the use of special equipment and these are the results broken down into a contingency table by age groups.

Notice that although age (in years) is a quantitative variable, the age group variable used in this analysis is categorical.

Let's begin with some simple probabilities, calculated logically using the relative frequency approach.

To aid our calculations, let's find the row, column, and overall totals.

10

Special Equipment?

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

- Select one person at random
- P(Oldest Age Group) = P(65+)
  = 3033 / 6582
  ≈ 0.4608
- P(Oldest Age Group **OR** Youngest Age Group)
  = (3033 + 792) / 6582 ≈ 0.5811 **Using logic**

  = P(Oldest) + P(Youngest)
  = (3033 / 6582) + (792 / 6582)
  ≈ 0.4608 + 0.1203
  ≈ 0.5811

  Using addition rule for disjoint events since oldest and youngest are disjoint. Use caution!

UF UNIVERSITY of FLORIDA

Here are the totals along with the original frequencies. If we select one person at random. What is the probability that the individual will be in the oldest age group?

This is a simple event where we only need to find the number of individuals in the age group and divide by the total number of individuals in our sample. Our answer is 3033 divided by 6582 or approximately 0.4608.

Now let's look at slightly more complex probability. Find the probability that the individual will be in the oldest or youngest age group. There were 792 in the youngest age group and 3033 in the oldest age group. There is no overlap, you cannot be in both of these age groups, so we haven't counted anyone twice.

This logical approach gives a probability calculated by adding 792 and 3033 and then dividing the result by the total, 6582 with a result of 0.5811.

We could apply probability rule #4, the addition rule FOR DISJOINT EVENTS, since these two events do not overlap, they are disjoint and we can simply add their probabilities together.

Using this approach uses the same numbers but in a slightly different way.

We find P(Oldest) = 3033/6582 which is approximately 0.4608 and then P(Youngest) = 792/6582 which is approximately 0.1203 and add these values together to get the same answer we found before, 0.5811.

This is useful for combining groups in the same categorical variable, let's look at a few more examples.

Special Equipment?

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

- P(45-64)
  = P(45-54 OR 55-64)
  = P(45-54) + P(55-64)
  = (1109 / 6582) + (1648 / 6582)
  ≈ 0.1685 + 0.2504
  ≈ 0.4189

Again – only because these are disjoint!

- P(45+)
  = P(45-54 OR 55-64 OR 65+)
  = P(45-54) + P(55-64) + P(65+)
  = (1109 / 6582) + (1648 / 6582) + (3033 / 6582)
  ≈ 0.1685 + 0.2504 + 0.4608
  ≈ 0.8797

UF UNIVERSITY of FLORIDA

To find P(45-64), we need to combine the results in the two age groups: 45-54 and 55-64. Using either logic or the probability rule #4 [the addition rule for disjoint events], we obtain 1109 out of 6582 for age group 45-54 and 1648 out of 6582 for age group 55-64 giving a probability of 0.4189.

When the events are disjoint, it is easy to extend this rule to as many events as needed – as long as they are all disjoint. Here for example, we can calculate the probability that a randomly selected individual will be 45 or older by combining the three age groups.

We need to find the probability that the individual is either in:
- age group 45-54 OR
- age group 55-64 OR
- age group 65 +.

An individual can only be in one of these groups and so they are disjoint and we can simply add their probabilities together as shown here to get 0.8797.

Do be careful to only apply this rule if you are completely certain that the events are disjoint.

Before we look at an example of "OR" where the events are NOT-disjoint using probability rule #5, we need to illustrate how to find the P(A and B) using logic.

Later, we will learn two rules for finding P(A and B), but these will mostly be used for finding probabilities in repeated sampling.

Notice in our next examples that we do not need ANY rule to find P(A and B). Be sure to avoid using a rule to find P(A and B) in similar problems.

One final comment about these questions. Although there are instances in practice where finding the P(A and B) or P(A or B) might be of interest, for the most part, the probabilities we will be finding in the next few examples have no clear practical use other than to illustrate these probability rules.

These probability rules are important but more in the underlying theoretical probability that arises in statistics than for the actual need to answer the kinds of questions we present.

When we get to the next section on conditional probability, that will change entirely and we will again have an excellent reason for understanding the interpretation of the results of our calculation in addition to gaining appreciation for the rules being presented.

Special Equipment?

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

- P(65+ and YES)
  = P(Both 65+ and YES)
  = (611 / 6582)
  ≈ 0.0928
- P(35-54 and NO)
  = (762 + 992) / 6582
  = 1754 / 6582
  ≈ 0.2665
- USING ONLY **LOGIC!**

To begin, let's find the probability that the randomly selected individual is in the oldest age group (65+) AND requires special equipment. In order to be counted, an individual must satisfy both of these events. There are 611 individuals who are in the 65+ age group and answered YES to the question about the need for special equipment.

We write P(65+ and YES) = 611 / 6582 ≈ 0.0928.

Note that we are still dividing by the total in this case.

As we learned in the exploratory data analysis unit for Case C-C, the result is much different when we restrict our calculation to a certain row or column. We will address those types of probabilities in the next section.

Let's look at one more. Find the P(35-54 and NO).

Now we have two age groups, but the result is similar. We ask, how many individuals are between the ages of 35-54 and also answered NO to the question about special equipment.

There are 762 individuals who answered NO in the 35-44 age group and 992 who answered NO in the 45-54 age group for a total of 762 + 992 [= 1754] dividing by 6582 gives an approximate probability of 0.2665.

These probabilities required no rule other than the relative frequency approach to calculating probability.

Now we move on to a few examples of P(A or B) when the events are not disjoint.

Special Equipment?

- From previous examples:
  - P(65+ and YES) ≈ 0.0928
  - P(65+) ≈ 0.4608
  - P(35-54 and NO) ≈ 0.2665
- P(65+ or YES) = ?
- P(35-54 or NO) ≈ ?

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

UF UNIVERSITY of FLORIDA

We will answer the corresponding "OR" probabilities for the last two examples.

We want to find P(65+ or YES) and P(35-54 or NO).

We can do this using logic or the general addition rule (rule #5).

We will complete each using both methods.

15

Special Equipment?

- Using logic:

P(65+ or YES)

= (978 + 2422) / 6582

= (3033 + (30 + 117 + 220)) / 6582

= (30 + 117 + 220 + 611 + 2422) / 6582

≈ 0.5166

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

UF UNIVERSITY of FLORIDA

Let's start with P(65+ or YES).

To determine the probability logically, we must count everyone who either answered YES or is 65+ without counting anyone twice. We need every frequency outlined in the 65+ row and YES column.

The easiest way would be to take the total for all those who answered YES which was 978 and then add to that the remaining individuals who were 65+ which would be 2422. This would give:

P(65+ or YES) = (978 + 2422) / 6582 = 0.5166

Alternatively we could take the total for those who were 65+ of 3033 and add to that the frequencies in the YES column for those in the other age categories (30 + 117 + 220). Or we could simply add all of the frequencies outlined in either orange or blue in the table, 30 + 117 + 220 + 611 + 2422.

Any logical way we approach the question should result in the same answer.

Notice that if we were to simply add 978 to 3033 we would have double counted the 611 individuals who are YES and 65+. Subtracting away the extra 611 individuals would solve the issue which is the approach taken by the general addition rule which we will illustrate next.

16

Special Equipment?

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

- From previous examples:
  - P(65+) ≈ 0.4608
  - P(65+ and YES) ≈ 0.0928
- P(YES) ≈ 0.1486
- Using Rule #5 – General Addition Rule

  P(65+ or YES)

  = P(65+) + P(YES) – P(65+ and YES)

  ≈ 0.4608 + 0.1486 – 0.0928

  ≈ 0.5166

UF UNIVERSITY of FLORIDA

To use probability rule #5, the general addition rule, to find P(A or B), we first need to have the three probabilities:

- P(A)
- P(B)
- P(A and B)

It is best to use only logic and counting to find these values. In this problem we need:

- P(65+)
- P(YES)
- P(65+ and YES)

We found the first and last probabilities before but let's go through them once again.

- There are 3033 individuals who are 65+ out of 6582 total. Dividing 3033 by 6582 gives P(65+) = 0.4608.
- There are 611 individuals who are both 65+ and YES out of 6582 total. Dividing 611 by 6582 gives P(65+and YES) = 0.0928.
- Now we need P(YES). There are 978 individuals who answered YES out of 6582 total. Dividing 978 by 6582 gives P(YES) = 0.1486.

Now we can simply use the rule #5 to get:

- P(65+ or YES)
- = P(65+) + P(YES) – P(65+ and YES)
- = 0.4608 + 0.1486 – 0.0928
- ≈ 0.5166

Which is the same value we found using logic alone.

Special Equipment?

- Using logic:

P(35-54 or NO)

= (5604+ (30+117)) / 6582

= ((792 + 1109) + (1428 + 2422)) / 6582

= (30 + 117 + 762 + 992 + 1428 + 2422) / 6582

≈ 0.8737

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

Now let's calculate P(35-54 or NO) using logic.

We need to count all outlined frequencies without counting anyone twice.

The easiest way to do this is to take the total for the NO column of 5604 and then add to that the frequencies in the YES column for individuals who are 35-44 and 45-54 which are 30 and 117 respectively.

We could also use the row totals for ages 35-44 and 45-49 (792 and 1109) and then add the remaining uncounted frequencies in the NO column (1428 and 2422).

Or finally, we could simply add all 6 numbers together 30, 117, 762, 992, 1428, and 2422.

Whichever method we use, we should find that P(35-54 or NO) is 0.8737.

18

Special Equipment?

Age:	Yes	No	Total
35-44	30	762	792
45-54	117	992	1109
55-64	220	1428	1648
65+	611	2422	3033
Total	978	5604	6582

- From previous examples:
  - P(35-54 and NO) ≈ 0.2665
- New values:
  - P(NO) = 5604 / 6582 ≈ 0.8514 (≈ 1 - 0.1486)
  - P(35-54) = (792 + 1109) / 6582 ≈ 0.2888
- Using Rule #5 – General Addition Rule

P(35-54 or NO) = P(35-54) + P(NO) – P(35-54 and NO)

≈ 0.2888 + 0.8514 – 0.2665

≈ 0.8737

UF UNIVERSITY of FLORIDA

To use probability rule #5, the general addition rule, to find P(A or B), we need:
- P(35-54)
- P(NO)
- P(35-54 and NO)

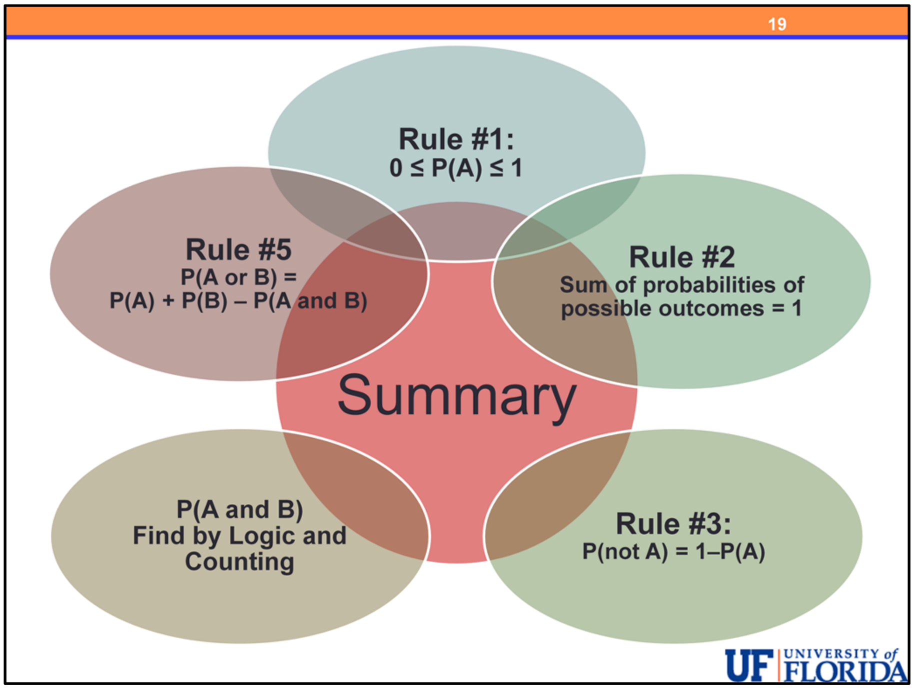
We found the last probability earlier but let's go through all of these calculations here
- There are 762 + 992 = 1754 individuals who are both 35-54 and NO out of 6582 total. Dividing 1754 by 6582 gives P(35-54 and NO) = 0.2665.
- There are 5604 individuals who answered NO out of 6582 total. Dividing 5604 by 6582 gives P(NO) = 0.8514. We could also have found this using P(YES) and the complement rule.
- Finally, there are 792 + 1109 = 1901 individuals who are 35-54 out of 6582 total. Dividing 1901 by 6582 gives P(35-54) = 0.2888.

Now we can simply use the rule #5 to get:
- P(35-54 or NO) = P(35-54) + P(NO) – P(35-54 and NO)
- = 0.2888 + 0.8514 – 0.2665
- ≈ 0.8737

Which is the same value we found using logic alone.

The most common mistake in applying this rule is in using a formula to find P(A and B).

It is difficult to make this point clear until we cover the rules for P(A and B) but note now that we only used logic and counting to find P(A and B) in the examples presented for finding P(A or B) here, in other videos, and in the course materials.



In the previous section we introduced basic concepts of probability including the relative frequency method of calculating probabilities from data and the law of large numbers which shows us that, in the long run, our estimates from representative data will approach the true value.

In this section, we presented some of the basic rules of probability and illustrated how to calculate basic probabilities from data using both logic and the rules.

We began with rules 1 and 2 which state some relatively simple rules of probability. That all probabilities are between 0 and 1 and that the sum of the probabilities of all possible outcomes must be equal to 1.

The complement rule, rule #3, states that the probability of the opposite of an event is 1 minus the probability of the event, i.e. P(not A) = 1 – P(A).

We defined disjoint events and discussed probability rule #4, the addition rule for disjoint events. Be very careful when you are applying a rule that has specific conditions such as this. You must be completely convinced that the conditions are satisfied in order to use such a rule. You will notice that this rule is not listed on this slide as I suggest using logic or starting from probability rule #5 since it works for any pair of events.

Then we illustrated how to find P(A and B) by logic and counting. And finally we discussed probability rule #5, the general addition rule for finding P(A or B). We illustrated using both logic and the rule to find P(A or B) from data.

BASIC PROBABILITY RULES

Unit 3A: Probability

UF UNIVERSITY of FLORIDA

You should notice that the types of examples we present in the course materials, videos, and quizzes are all very similar. You can expect to see questions like these again!