BASIC PROBABILITY

Yesterday we introduced our basic probability model. In this handout, we review the concepts involved in that model and use the model properties to develop more properties of probability. We also explore some applications to public health.

Recall Terminology:

Event: Something that might happen, with some degree of uncertainty.
Union of events: The event that one or more of the events happens.
Mutually exclusive events: If any one happens, then the others cannot.

Recall Basic Properties:

Property 1: $0 \leq P(E) \leq 1$
Property 2: P(S) = ___ (S = certain event; sample space)
Property 3: P(union of mutually exclusive events) = ________________________

More Terminology: The *intersection* E∩F (or EF for short) of two events is the event that *both* E and F are (simultaneously) the case.

- What is the connection between the concepts “intersection” and “mutually exclusive”?

Example: Suppose that the only people we are considering are the 116 people in the popcorn production plant study. Here is the two-way table from that problem:

	Low exposure	High exposure	Total
Airway obstructed	6	15	21
Airway not obstructed	52	43	95
Total	58	58	116

Let A be the event “Airway obstructed”. Let L be the event “Low exposure.” Describe the event A∩L: ____

Since we are only considering these 116 people, the probability of an event is the probability that a person randomly selected from the group studied was in the group describing that event – that is, the frequency of occurrence of that event. So

P(A) = ____ P(L) = ____ P(A∩L) = ____

Still More Terminology: The *complement* E^c of an event E is the event that E does not happen/is not the case.

In the example above:

i. Describe:

A^c : ______________________ L^c: ____________________

$A^c \cap L$: ____

ii. Identify two mutually exclusive events and check that their probabilities add as required.

More Properties of Probabilities:

i. E and E^c can't both happen at the same time, so E and E^c are ______________________.

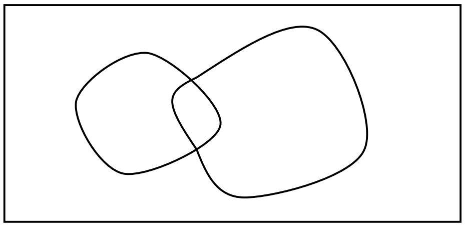
ii. $E \cup E^c$ = ________________________________

From (i), (ii), and Property 3 we conclude:
$P(E) + P(E^c)$ = _________ = ___,

so $P(E^c)$ = ______________ (***Property 4***)

Check this out in the Example with E = A:

Things are now getting a little more complicated, so it can help to draw pictures (called "Venn Diagrams"—Venn in doubt, draw a diagram!) like this:



- Label the left shape E, the right shape F.
- Identify each of the following in the diagram: $E \cap F$, $E \cup F$, E^c, F^c, EF^c, and $(EF)^c$.
- Use the diagram to convince yourself that you need to be careful not to confuse EF^c (that is, $E \cap F^c$) and $(EF)^c$ (that is, $(E \cap F)^c$).

iii. $(EF) \cap (EF^c)$ = _______ (A Venn diagram can help), so EF and EF^c are __________________________________.

iv. $EF \cup EF^c$ = ____ (A Venn diagram can help).

From (iii), (iv), and Property 3, we conclude:
$P(E)$ = __________+_____________. (***Property 5***)
Similarly,
$P(F)$ =
Check these out in the Example with A and L:
Note: Property 5 can be rearranged to give us two more "identities":

$P(EF)$ =
$P(EF^c)$ =

v. (Look at the Venn diagram) $E \cup F = EF \cup EF^c \cup$______

vi. EF, EF^c, and _______ are mutually exclusive.

From (v), (vi), and Property 3, we conclude:

$P(E \cup F) =$ __________________________ (***Property 6***)

Check this out in the Example with A and L:

Note: Combining Property 6 with Property 5, we get

P(EUF) = P(E) + P(F) - _______

(Fill in details of how!)

Conditional Probability

P(E|F) ("The probability of E given F") is the probability that event E occurs, assuming that event F occurs.

Note There is no implication of causality involved; F does not have to occur before E.

Example: Calculate directly from the data:

P(Low Exposure | Airway obstructed) =

P(Airway obstructed| Low exposure) =

So P(E|F) *and* P(F|E) *are different!* (i.e., same concept, different applications)

This example motivates (How?) the following **formal definition of conditional probability**:

$P(E|F) = P(EF)/P(F)$

Note: This definition assumes that $P(F) \neq 0$. (and hence that $P(F) > 0$.)

Note: This definition can be thought of as "rescaling probabilities given the event F to make total probability = 1". This is legitimate because *P(-|F) is itself a probability function*. To see this, we need to prove that P(-|F) satisfies properties (1) - (3) of a probability function.

Proof of Property 1 for P(- |F): By Property 1 for P, $P(E \cap F) \geq 0$. It follows from this $P(E \cap F)/P(F) > 0$ – in other words, $P(E|F) > 0$.

Proof of Property 2 for P(- |F): Recall that the sample space is called S. So we need to show that $P(S|F) = 1$. This follows from the formal definition:

$$P(S|F) = P(S \cap F)/P(F) = P(F)/P(F) = 1.$$

The idea of the proof of Property 3 for P(- |F): For notational convenience, we will only show this in the case of two mutually exclusive events E_1 and E_2. Since these events are mutually exclusive, Property (3) for the original probability function P( )tells us that

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

To have property 3 for P(- |F), we need to show

$$P((E_1 \cup E_2)|F) = P(E_1|F) + P(E_2|F)$$

Now

$$P((E_1 \cup E_2)|F) = P((E_1 \cup E_2) \cap F)/P(F) \qquad \text{(Why?)}$$

A Venn diagram will show that $(E_1 \cup E_2) \cap F = (E_1 \cap F) \cup (E_2 \cap F)$. Also note that since E_1 and E_2 are mutually exclusive, $E_1 \cap F$ and $E_2 \cap F$ are also mutually exclusive. Using both of these properties,

$$\begin{aligned} P((E_1 \cup E_2)|F) &= [P((E_1 \cap F) \cup (E_2 \cap F))]/P(F) \\ &= [P(E_1 \cap F) + P(E_2 \cap F)]/P(F) \\ &= P(E_1 \cap F)/P(F) + P(E_2 \cap F)]/P(F) = P(E_1|F) + P(E_2|F) \end{aligned}$$

Note: One consequence of the fact that P(- |F) is a probability function is that particular,

$$P(E^C|F) = 1 - P(E|F).$$

We can also check this out directly as follows: Apply Property 5 with the roles of E and F reversed to get

$$P(F) = P(EF) + P(E^cF) \qquad \text{(Draw a Venn diagram.)}$$

Divide by P(F)and use the definition of conditional probability to get

$$1 = P(E|F) + P(E^c|F)$$

vi. From the definition of conditional probability, we conclude:

$P(E \cap F) =$ ___________ (***Property** 7*)

and symmetrically,

$P(E \cap F) =$ ______________ (***Property 7 in another form***)

Check this out in the example with A and L.

Comment: Often when we talk about probabilities, we are talking about conditional probabilities without explicitly using the notation. For example, in the examples above, we said, "Suppose that the only people we are considering are the 116 people in the popcorn production plant study." Instead we could have written:

P(Low exposure | in the study) instead of P(Low exposure)
P(Airway obstructed | in the study) instead of P(Airway obstructed)
P(Low exposure | airway obstructed and in the study) instead of
P(Low exposure | airway obstructed), etc.

Similarly, in the Lipitor study mentioned in Problem 3c of Part II of the handout *Measures, Words, Rates, Ratios, and Proportions*, we could say that
P(stroke| has type 2 diabetes and one other risk factor for heart disease and takes Lipitor) was estimated to be .015.

Exercises: Not explicitly stating the conditions can lead to confusion. Here are some examples.

1. Many people aren't clear what "The probability that it will rain tomorrow is 40%" means. Which of these do you think it means?

a. It will rain 40% of the time tomorrow.
b. Tomorrow it will rain in 40% of the area covered by the forecast.
c. On 40% of days with conditions like today, the next day will have some rain.
d. None of the above.
e. No idea

2. A physician tells a patient, "If you take Prozac, you have a 30% to 50% chance of developing a sexual problem, such as impotence or loss of interest." Which of the following do you think this means? Might the patient interpret the statement differently from the physician?

a. The patient will have such a problem in 30 to 50% of the occasions in which he might otherwise have sex.
b. 30 to 50% of patients taking Prozac develop such problems.
c. None of the above.
d. I'm clueless

Note:

- Notice that the specific alternatives given above all refer to what would be in the *denominator* of the probability calculation.
- Some people refer to the cause of confusion in the above examples as "not explicitly stating the reference class."
- This confusion often occurs in statistical inference; more on this later.

Problems:

→ *Before you start these problems, make yourself a list (for handy reference) of the properties of probability that we have established so far.*

1. Using the properties of probabilities that have been discussed so far, prove the following additional property: $P(E) = P(E|F)P(F)+P(E|F^c)P(F^c)$. *Do this starting with the left hand side of the equation and working toward the right hand side. Be sure to give the reason (i.e., which property you have used) for each step.*

2. Medical tests for diseases give results that are called "positive" and "negative." For an ideal test, the result would be positive if and only if the patient has the disease, and negative if and only if the patient does not have the disease. Unfortunately, there are no ideal tests, so we need to consider *false positives* (when the test result is positive but the patient does not have the disease) and *false negatives* (when the test result is negative but the patient does have the disease). Hence we need to consider the following two probabilities:

The *false positive probability* (or *false positive rate)*:

P(positive test result | disease not present)

The *false negative probability* (or *false negative rate)*:

P(negative test result |disease present))

The false positive and false negative rates for a medical test are usually determined before the test is widely used.

Suppose that a certain test has false positive rate 0.05 and false negative rate 0.01, and that 2% of the population has the disease. Use the result in Problem 1 to calculate the probability of having a positive result on the test. [*Hint*: Let E be the event "Test is positive" and F the event "Has the disease". You will need to figure out from the given information and the properties of probabilities what P(test is positive | disease present) and P(does not have the disease) are.]

3. Using the properties of probabilities that have been discussed so far, prove the following additional property ("Bayes' rule"):

$P(F|E) = P(E|F)P(F)/P(E)$.

As in Problem 1, do this starting with the left hand side of the equation and working toward the right hand side. Be sure to give the reason (i.e., which property you have used) for each step.

(More problems on next page)

4. a. In the situation of Problem 2, use Bayes' rule (from Problem 3) and the result of Problem 2 to find P(has the disease | test is positive), the probability of having the disease if you test positive. *Note*: This probability is called the *positive predictive value.* (If you get stuck, try part (b) and then come back to part (a).)

b. Do the problem in part (a) (that is, calculate P(has the disease | test is positive) another way, and check with your answer to part (a). One way is to call the total population size T and make a table like the one below, then figure out missing parts to do the needed calculation:

	Has disease	Does not have disease	Total
Test is positive			
Test is negative			
Total			T

If you don't get the same answer both ways, check over each method to see if you can find errors.

c. Imagine you are talking to your grandmother, uncle, or neighbor who has only an eighth grade education. They tell you that their physician gave them the test and told them that they tested positive, that 5% of people who don't have the disease test positive, 1% of people who have the disease test negative, and 2% of people their age and gender have the disease. (This is just the information given above, but phrased differently.) What do you think would be a good way to explain to them what the chances are that they have the disease?

5. a. How would your answer to problem 4 be different if 10% of the population had the disease (instead of 2%)? If only 0.1% of the population had the disease? (Assume the same false positive and false negative rates.)

b. Find a formula for the positive predictive value (i.e., P(has the disease | test is positive)) in terms of the *prevalence rate* r (that is, the percent of people in the population who have the disease), and/or set up an Excel spreadsheet to do this calculation.

6. Derive a generalization of Property 7 to several events. That is, derive a formula for $P(E_1E_2 \ldots E_n)$ in terms of conditional probabilities, as a product ending in $P(E_{n-1} \mid E_n)P(E_n)$. [Hint: Try it first with just three events.] Be sure to give reasons for every step.