340.721 Epidemiologic Inference in Public Health I

PRE-Activity Questions:
Outbreak Investigation

The Activities provide experience in applying epidemiologic methods, interpreting findings, and drawing inferences. Activities will be discussed during the LiveTalks. Students are expected to work with their assigned Course Group prior to the start of each LiveTalk.

Prior to each Activity, students are to complete the corresponding set of PRE-Activity Questions. Each set of PRE-Activity Questions consists of 10 graded multiple choice questions. The graded multiple choice questions are to be completed via CoursePlus by the date and time listed in CoursePlus. PRE-Activity Questions prepare you for a productive and collaborative experience during the Activities.

Expectations for the PRE-Activity Questions

1. *Individually, read and attempt to answer all PRE-Activity Questions.*

2. *"Meet" or communicate with fellow students discuss challenging concepts, questions and compare answers. You may refer to their course materials and are strongly encouraged to collaborate with fellow students to complete the PRE-Activity Questions.*

3. *PRE-Activity Questions are due to Courseplus by the date listed on the syllabus. Although group collaboration is encouraged to complete the PRE-Activity Questions, each student must individually submit the PRE-Activity Questions.* ***Without exception, no credit will be given for submitting the PRE-Activity Questions after the due date****. The lowest PRE-Activity grade will be dropped when calculating the overall course grade.*

Motivation
Epidemiologists are often involved in outbreak investigations to try to identify the cause(s) of the outbreak in order to prevent another outbreak. Consider the following motivating questions for this assignment: Is there an epidemic?; How do you define a case?; Are the number of cases normal or above normal?; If there is an epidemic what is causing it?; Who is at risk?; What is the mode of transmission?; What is the timing of the epidemic?; Is this timing similar to other known causes?; Is the cause contained?; If not what changes do you recommend to contain the cause?

This assignment corresponds to:
Lectures: Epidemics
Readings: Gordis text (5^{th} ed.) Chapter 2

Introduction
This assignment describes an outbreak of sore throat that took place in Baltimore some years ago and was investigated by the Baltimore City Health Department. A classic JHSPH Department of Epidemiology exercise, this assignment will give you an idea of how such an investigation is carried out. In addition, it will introduce you to a number of important concepts, including the epidemic curve, the median incubation period, the attack rate and the relative risk. It will also demonstrate how an investigator attempts to assess possible causal factors in order to identify the one most likely to be the cause of the disease.

Concepts covered:
1. Definition of a case
2. Selection of a study population
3. Epidemic curve
4. Incubation period
5. Attack rate
6. Relative risk
7. Attributable risk (risk difference)
8. Cross-tabulation

Learning Objectives:

1. Utilize principles for case definition
2. Identify selection criteria for a study population
3. Draw an epidemic curve
4. Determine an incubation period
5. Calculate attack rate
6. Calculate relative risk
7. Calculate attributable risk (risk difference)
8. Compose a cross-tabulation

NOTE: *For students who would like an introduction to outbreak investigation, including the calculation of median incubation period, please see Problem 0 in Courselpus (answer key included).*

STEPS IN OUTBREAK INVESTIGATION:

1. Define the epidemic (case definition, population at risk, attack rates).
2. Examine the distribution of cases by time and place (epidemic curves, median incubation period)
3. Look for combinations of relevant variables (food-specific attack rates, cross-tabulation).
4. Develop hypotheses.
5. Test hypotheses
6. Recommend control measures

DESCRIPTION OF THE OUTBREAK

On Friday, February 8, a private physician telephoned the Baltimore City Health Department to report that on the day before and the morning of his call, he had seen several women with acute sore throat. Each of the women gave a history of attendance at a luncheon at a National Guard Armory on Wednesday, February 6 at 12:00 noon. The physician described the illnesses as characterized by acute onset with chills, fever, general malaise, and sore throat; physical examination revealed inflamed throat with some exudate, cervical adenitis, and temperature between 102-104 F.

In addition, he stated that one of the husbands, who had not attended the luncheon, had an acute sore throat. The physician further observed that the wife had brought home some leftover food and her husband had eaten this for supper on February 6.

An investigation was undertaken immediately, and the District Health Officer, who made the first home visits, verified the physician's original report and obtained additional information that indicated that this was an outbreak of major proportions.

The luncheon had been a fund-raising effort to help fight cancer in children and had been an annual affair of an organization of 96 women for several years. Between 800-900 people, mostly women, attended the luncheon.

The women had largely conducted the procurement and preparation of the food that was served, although a commercial caterer and a restaurateur had voluntarily assisted in the preparation of the food. The complete menu, which was served cafeteria style, was as follows: Egg Salad, Tuna Fish Salad, Macaroni and Cheese, Cottage Cheese with Nuts and Cherries, Pickles and Olives, Ice Cream, Coffee and Cookies.

A questionnaire was prepared by the Health Department and distributed one week after the luncheon to as many people who had attended as possible. The questionnaire requested the following information: clinical details of the illness; time of onset of the illness; name of the attending physician; history of foods consumed at the luncheon; a statement whether any food had been taken home and if so, who had eaten it with what result. The notes below summarize the data collected. If you would like to see the raw data for all 96 members of the organization and the 67 guests who responded to the questionnaire, an Excel file of these data is available on the course website.

NOTES SUMMARIZING THE DATA FROM THE OUTBREAK:

Population:	Members who attended the luncheon:	90
	Guests who attended the luncheon:	~800

Symptoms (Members only):

Any Symptom:	61
Sore Throat:	60
Fever:	43
Headache:	33
Vomiting:	5
Diarrhea:	8

Symptoms (Guests only):

Any Symptom:	58
Sore Throat:	57
Fever:	49
Headache:	32
Vomiting:	8
Diarrhea:	11

DEFINE THE EPIDEMIC

CASE DEFINITION

First, consider how to <u>define a case</u> of illness. How would you decide which symptom(s) you would include in your case definition? There are several possibilities and each definition will yield a different number of cases.

A more strict definition would require more symptoms to be included in the definition of a case. For example, a case might be defined as people who report sore throat AND fever AND vomiting. Conversely, a less strict definition would require fewer symptoms (for example, only headaches). Each definition will differ in its ability to properly classify those individuals who are truly sick and those individuals who are truly <u>*not*</u> sick. For example, when using a *more* strict case definition, fewer individuals will be identified as cases. This means that you are likely to miss some cases among those individuals who were truly ill, but also that more individuals who truly are <u>*not*</u> ill will be correctly counted as non-cases. The case definition should match the goal of the investigation. In this Assignment, the goal of the investigation is to include as many "true" cases and as few "false" cases as possible.

<u>Question 1</u>
Which of the following statements is(are) true of using a **less strict** case definition as compared to a more strict case definition? **(SELECT ALL THAT APPLY)**

a. More individuals who truly are ill will be counted as cases
b. More individuals who truly are ill will be counted as non-cases
c. More individuals who truly are <u>*not*</u> ill will be counted as non-cases
d. More individuals who truly are <u>*not*</u> ill will be counted as cases

DESCRIBE THE STUDY POPULATION AT RISK

In addition to identifying a case definition, you must decide who to include in your study. Ideally, you would obtain measurements on everyone who was at risk of becoming ill.

Question 2

The term ‘at risk’ is defined as those individuals who are able become ill in the same way. Which of the following groups can be considered ‘at risk’ in this outbreak if we believe the outcomes are due to a food-related illness?

a. All members and their families
b. All members and guests who attended the luncheon
c. All members and guests who attended the luncheon and ate
d. All members and guests who attended the luncheon, ate, and reported having symptoms

Recall that a questionnaire was prepared by the Health Department and distributed one week after the luncheon to all members and to as many guests as possible in order to obtain clinical details of the illness. Table 1 summarizes the responses from members and from guests.

Table 1. Summary of Response to the Questionnaire for Members and Guests

	Members	**Guests**
Number of people to whom questionnaires were sent	90	As many as possible (Assume ~800)
Number of responses from attendees	86	67
Number of responses from attendees who ate	85	67
Number of people who reported sore throat	60	57
Number of people who reported sore throat and fever	41	48

From Table 1, 90 members were asked to complete the questionnaire. 86 of those 90 members returned the questionnaire. The questionnaire response rate for members is therefore:

$$\text{Response rate} = \frac{\text{\# of members that completed the questionnaire}}{\text{\# of members who were asked to complete the questionnaire}} = \frac{86}{90} = 0.96, \text{or } 96\%$$

Question 3
What is the response rate for guests?

a. 8%
b. 46%
c. 72%
d. 85%

(NOTE: Are you wondering why you were asked to calculate the response rate? Compare the response rate you calculated for guests in Question 3 to the 96% response for members. Are they similar or different? What are some possible reasons why?

If you are unable to obtain measurements on everyone, a more practical goal is to obtain measurements from *a representative group* of all the people at risk of becoming ill. Based on the response rates, which group do you think is more representative – members or guests?)

ATTACK RATES

The *attack rate* is an example of a measure of disease frequency. It is calculated as the number of people at risk in whom a certain illness develops divided by the total number of people *at risk*. Using a case definition of sore throat and a study population that includes only 'at risk' members the overall Attack Rate in members is:

$$Attack\ rate = \frac{= \text{ \# of members who attended the luncheon and ate and who then reported sore throat}}{\text{\# of members who attended the luncheon and ate}}$$
$$= \frac{60}{85} = 0.706, or\ 70.6\%$$

Attack rates can help to identify the probable source of an outbreak. Food-specific attack rates can determine whether a particular food may be implicated as the source of the agent causing the illness. Table 2a summarizes the total number of members who ate and did not eat a specific food, and the number of members who ate that food and reported becoming ill.

The food-specific attack rate is calculated as the number of people at risk who ate a specific food in whom a certain illness develops divided by the total number of people at risk who ate that specific food.

Using a case definition of sore through, the food-specific attack rate for people who ate egg salad is:

$$Attack\ rate_{egg\ salad} = \frac{54}{65} = 0.831\ or\ 83\%$$

The attack rates for a case definition of sore throat have been provided for you in Table 2a. Complete Table 2a by calculating the attack rates for a case definition of sore throat and fever.

Table 2a. Attack Rates for Members Who Ate and Did Not Eat a Specific Food for the Case Definition of Sore Throat and the Case Definition of Sore Throat & Fever

Food	**Ate the Specified Food**			**Did Not Eat the Specified Food**		
	Total	**Became Sick**	**Attack Rate (%)**	**Total**	**Became Sick**	**Attack Rate (%)**
CASE DEFINTION: SORE THROAT						
Egg Salad	65	54	83	20	6	30
Macaroni & cheese	34	26	76	51	34	67
Cottage Cheese	56	40	71	29	20	69
Tuna Salad	63	49	78	22	11	50
Ice Cream	40	31	78	45	29	64
Other	79	57	72	6	3	50
CASE DEFINITION: SORETHROAT & FEVER						
Egg Salad	65	38		20	3	
Macaroni & cheese	34	20		51	21	
Cottage Cheese	56	29		29	12	
Tuna Salad	63	36		22	5	
Ice Cream	40	21		45	20	
Other	79	41		6	1	

Question 4

Using the data from Table 2 and a case-definition of sore throat and fever, what is the food-specific attack rate for people who ate **tuna salad**? (Recall that this study population only includes 'at risk' members.)

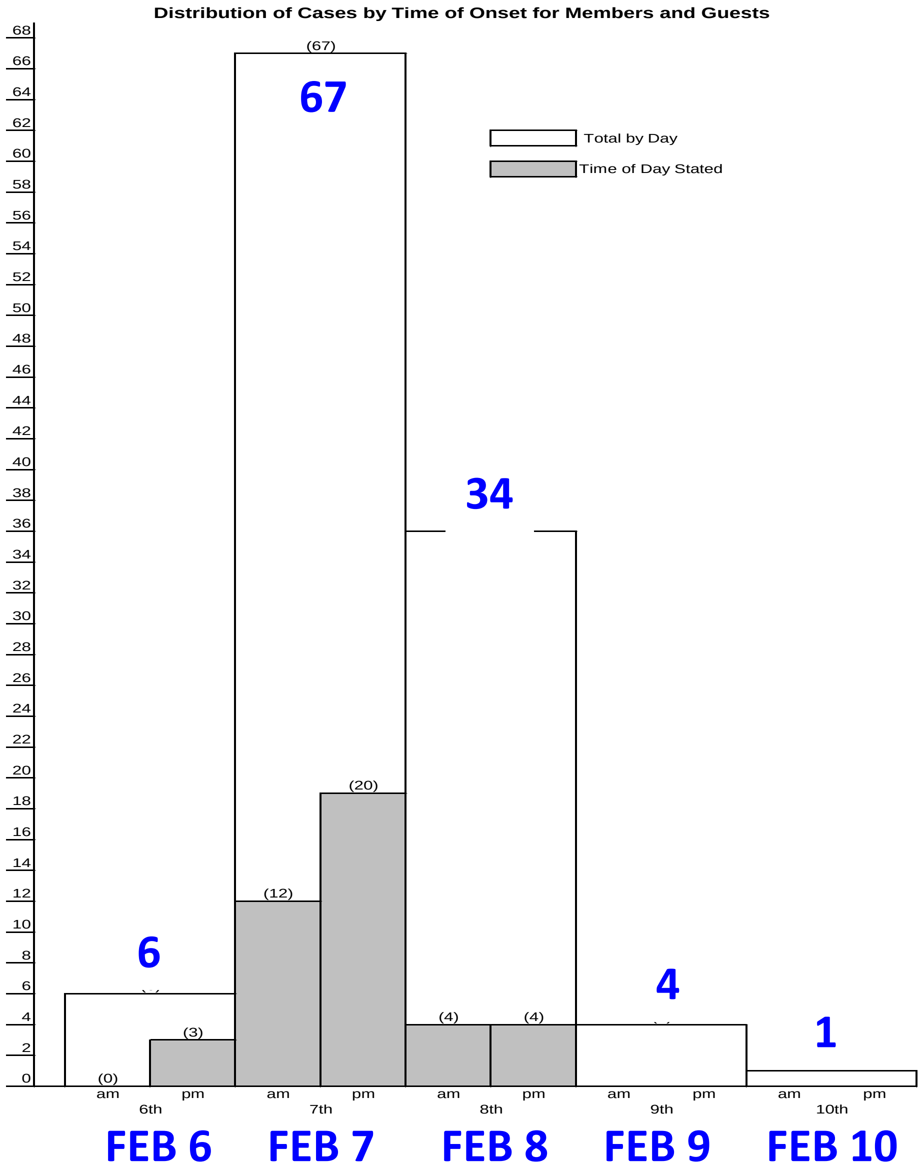
a. 23%
b. 35%
c. 57%
d. 78%

EXAMINING THE DISTRIBUTION OF CASES BY TIME AND PLACE

EPIDEMIC CURVE

Figure 1 shows the graph of the distribution of cases by day of onset, and where possible by time of day. This plot is the **epidemic curve**. You can see that the largest number of cases falls on February 7.

Figure 1. Distribution of Cases by Time of Onset for Members and Guests



If you wish, you can re-create this plot using the data in the Excel spreadsheet on Courseplus. At the end of this assignment, we will practice plotting an epidemic curve using data from a smaller outbreak.

POSSIBLE ETIOLOGIC AGENT

Table 4 summarizes the incubations period, symptoms and modes of transmission for several possible agents that may have caused the outbreak. The incubation period for illness in this outbreak is ~28 hours. (You can calculate this incubation period using the data provided on Courseplus and by making a couple of assumptions – ask a TA about it! You will practice calculating the median incubation period using data from a smaller outbreak at the end of this assignment.)

Table 4. Possible etiologic agents by incubation period, symptoms and mode of transmission

Etiologic agent	**Incubation period**	**Symptoms**	**Transmission**
Salmonella enterica	6 to 72 hours	Sudden onset Headache Abdominal pain Diarrhea Nausea	Ingestion of contaminated food, raw and undercooked eggs, raw milk Fecal-oral transmission
Streptococcus group A	1 to 3 days	Sudden onset Fever Sore throat Swollen lymph nodes	Droplet inhalation Direct contact with an infected person Ingestion of contaminated food
Corynebacterium diphteriae	2 to 5 days	Sore throat Swollen lymph nodes Grayish lesions in the throat	Direct contact with an infected person Contact with articles soiled with discharges from lesions Ingestion of raw milk
Coxsackievirus group A	3 to 5 days	Sudden onset Fever Sore throat Grayish lesions in the throat	Droplet inhalation Direct contact with nose and throat discharges or feces of infected person
Escherichia coli O157:H7	2 to 8 days	Diarrhea	Ingestion of contaminated food, inadequately cooked beef, raw milk
Mycoplasma pneumonae	6 to 32 days	Headache Malaise Cough Sore throat	Droplet inhalation Direct contact with an infected person

Chin J, ed. *Control of Communicable Diseases Manual*, 17^{th} Edition. APHA: 2000.

Question 5

Based on the information provided earlier in the assignment and in Table 4, what is the etiologic agent that is most likely responsible for the outbreak?

a. Streptococcus group A
b. *Corynebacterium diphteriae*
c. Coxsackievirus group A
d. Escherichia coli O157:H7

COMBINATIONS OF RELEVANT VARIABLES

COMPARING FOOD-SPECIFIC ATTACK RATES

In Question 4, you calculated a food-specific attack rate for tuna salad. Yet to complete Table 2, you also calculated the attack rate in people who had the same case definition, but did *not* eat tuna salad. Why? To compare attack rates in people who ate a specific food to people who did not eat that food. (Which attack rate would you expect to be larger if the food is implicated as the source of the agent causing the illness in the outbreak?)

We can compare attack rates in people who ate and did not eat a specific food by taking a difference of the attack rates, or by calculating a ratio of the two attack rates.

For example, the difference in attack rates in people who did and did not eat egg salad for a case definition of sore throat is: 83% - 30% = 53%

The ratio of the two attack rates is: $\frac{83\%}{30\%} = 2.8$

The attack rate differences and attack rate ratios for a case definition of sore throat have been provided for you in Table 2b. Complete Table 2b by calculating the attack rate difference and attack rate ratios for a case definition of sore throat and fever.

Table 2b. Attack Rates Differences and Attack Rate Ratios for Members Who Ate and Did Not Eat a Specific Food for the Case Definition of Sore Throat and the Case Definition of Sore Throat & Fever

Food	Ate the Specified Food			Did Not Eat the Specified Food			Difference in Attack Rates (%)	Ratio of Attack Rates
	Total	Became Sick	Attack Rate (%)	Total	Became Sick	Attack Rate (%)		
CASE DEFINTION: SORE THROAT								
Egg Salad	65	54	83	20	6	30	53	2.8
Macaroni & cheese	34	26	76	51	34	67	9	1.3
Cottage Cheese	56	40	71	29	20	69	2	1.0
Tuna Salad	63	49	78	22	11	50	28	1.6
Ice Cream	40	31	78	45	29	64	14	1.2
Other	79	57	72	6	3	50	22	1.4
CASE DEFINITION: SORETHROAT & FEVER								
Egg Salad	65	38		20	3			
Macaroni & cheese	34	20		51	21			
Cottage Cheese	56	29		29	12			
Tuna Salad	63	36		22	5			
Ice Cream	40	21		45	20			
Other	79	41		6	1			

Question 6

For a case definition of sore throat and fever, what is the *difference* between the two attack rates and the *ratio* of the two attack rates for people who did and did not eat egg salad?

a. Difference: 23%; Ratio: 2.5
b. Difference: 43%; Ratio: 3.9
c. Difference: 58%; Ratio: 1.5
d. Difference: 83%; Ratio: 2.8

CROSS-TABULATION

Given the results in **Table 2b**, egg salad and tuna salad are both implicated in the outbreak. But people who eat egg salad may also have eaten tuna salad and vice versa which makes it difficult to determine which food was responsible for the outbreak. We can calculate separate estimates of the risk associated with each food using a cross-tabulation. The goal of the cross-tabulation is to compare attack rates in persons who ate a given food *independent* of other foods in order to determine which food was likely the cause of the outbreak.

First, for a case definition of *sore throat*, let's look at the number of members who were sick and were well among the members who ate only tuna salad, ate only egg salad, ate both tuna and egg salad and ate neither food:

Table 3a. Cross-tabulation Table for Members with Sore Throat

	Number Sick	**Number Well**	**Total**	**Attack Rate**
Ate Tuna Salad only	3	7	10	0.30
Ate Egg Salad only	8	4	12	0.67
Ate both Tuna and Egg Salad	46	7	53	0.87
Didn't eat either	3	7	10	0.30

Let's organize the attack rates in a different way:

Table 3b. Cross-tabulation Table of Attack Rates for Members with Sore Throat

		Ate Egg Salad	
		Yes	**No**
Ate Tuna Salad	**Yes**	0.87	0.30
	No	0.67	0.30

The attack rate in people who ate only tuna salad (i.e., did not eat egg salad) is 0.30 or 30%.

The attack rate in people who ate only egg salad (i.e., did not eat tuna salad) is 0.67, or 67%.

The attack rate in people who ate neither food is 0.30 or 30%. *This attack rate gives us an estimate of the background risk of disease in this population*.

To calculate the excess risk over and above the background risk in the population that is associated with eating tuna salad, we can divide the attack rate in people who only ate tuna salad by the attack rate in people who ate neither food:

$$\frac{Attack\ rate\ in\ people\ who\ ate\ only\ tuna\ salad}{Attack\ rate\ in\ people\ who\ ate\ neither\ food} = \frac{0.30}{0.30} = 1.0$$

This number is called a *relative risk*. The null value for a relative risk (i.e., no increased risk compared to the background risk) is 1. We therefore conclude that there is no excess risk of illness associated with eating tuna salad.

Similarly, the ratio of the attack rate in people who ate only egg salad by the attack rate in people who ate neither food is:

$$\frac{Attack\ rate\ in\ people\ who\ ate\ only\ egg\ salad}{Attack\ rate\ in\ people\ who\ ate\ neither\ food} = \frac{0.67}{0.30} = 2.2$$

Because this ratio is greater than 1, we conclude that the egg salad was the cause of the outbreak.

Complete Table 4a and compose a cross-tabulation (Table 4b) by calculating the attack rates for members who ate only tuna salad, ate only egg salad, ate both tuna and egg salad and ate neither food for a case definition of *sore throat and fever*.

Table 4a. Cross-tabulation Table for Members with Sore Throat & Fever

	Number Sick	**Number Well**	**Total**	**Attack Rate**
Ate Tuna Salad only	2	8	10	
Ate Egg Salad only	4	8	12	
Ate both Tuna and Egg Salad	34	19	53	
Didn't eat either	1	9	10	

Table 4b. Cross-tabulation Table of Attack Rates for Members with Sore Throat

		Ate Egg Salad	
		Yes	**No**
Ate Tuna Salad	**Yes**		
	No		

<u>Question 7</u>

What is the relative risk of eating only egg salad and eating only tuna salad compared to the background risk of disease for a case definition of sore throat and fever?

a. Egg salad: 1.0; Tuna salad: 2.0
b. Egg salad: 2.0; Tuna salad: 3.3
c. Egg salad: 3.3; Tuna salad: 2.0
d. Egg salad: 3.3; Tuna salad: 6.4

What do you conclude about the cause of the outbreak?

EPIDEMIC CURVE AND MEDIAN INCUBATION PERIOD

On September 12, 20 people attended an office party where lunch was served at 1:00 pm. An outbreak of vomiting and diarrhea occurred following the party and is thought to be due to either the cake or the cookies that were served with lunch. A summary of the health status, illness onset, and the foods eaten for the 20 attendees is shown in the table below.

Person Number	**Health Status**	**Time of Onset of Symptoms**	**Foods Eaten**
1	Sick	September 12th, 5:30 P.M.	Cake only
2	Sick	September 12th, 5:30 P.M.	Cookies only
3	Sick	September 12th, 6:30 P.M.	Both cake and cookies
4	Sick	September 12th, 6:30 P.M.	Neither cake nor cookies
5	Sick	September 12th, 6:30 P.M.	Both cake and cookies
6	Sick	September 12th, 8:00 P.M.	Cake only
7	Sick	September 12th, 8:00 P.M.	Cookies only
8	Sick	September 12th, 8:30 P.M.	Cookies only
9	Sick	September 12th, 9:30 P.M.	Cookies only
10	Sick	September 12th, 10:30 P.M.	Both cake and cookies
11	Sick	September 12th, 11:00 P.M.	Both cake and cookies
12	Sick	September 13th, 12:30 A.M.	Cake only
13	Sick	September 13th, 2:30 A.M.	Cake only
14	Well		Neither cake nor cookies
15	Well		Both cake and cookies
16	Well		Neither cake nor cookies
17	Well		Cookies only
18	Well		Neither cake nor cookies
19	Well		Neither cake nor cookies
20	Well		Cake only

EPIDEMIC CURVE

Question 8

Plot the epidemic curve for this outbreak. Which time of day on September 12 has the largest number of cases?

a. 5:30 PM
b. 6:30 PM
c. 8:00 PM
d. 10:00 PM

MEDIAN INCUBATION PERIOD

(NOTE: See Problem 0 for practice plotting an epidemic curve and calculating the median incubation period)

Question 9

Using the epidemic curve you plotted in Question 8, what was the median incubation period for the onset of vomiting and diarrhea?

a. 4 hours, 30 minutes
b. 7 hours
c. 10 hours, 30 minutes
d. 13 hours, 30 minutes

Question 10

Using the cross-tabulation method you learned in Question 7, what is(are) the most likely contaminated food(s)?

a. Cake only
b. Cookies only
c. Both cake and cookies
d. Neither cake nor cookies