

One answer and only one answer per question. Leaving a question blank or filling in 2+ answers will be incorrect no matter what.

Where relevant, the goal is underlined. *Italicized phrases are true*. Do not assume more than is given in a question.

A = True, B = False unless indicated otherwise. If any part of an answer is incorrect, treat all of it as incorrect. If different parts of an option are inconsistent with each other, consider it incorrect.

Data Framing

1-4 (3 pts) In a histogram figure (a figure with vertical bars), what is the effect of truncating the vertical axis so that it starts at a value larger than zero? Suppose, for example, that the data consist of values 1000 for one bar and 1500 for the other bar. What is the effect of starting the vertical axis at 500 instead of 0? You may rely on examples shown in class. **A = True, B = False**

1. **(A)(B)** The truncation makes the visual appearance of the difference between the two numbers appear larger than it is.
2. **(A)(B)** The truncation makes the visual appearance of the difference between the two numbers appear smaller than it is.
3. **(A)(B)** The truncation alters the the data shown – the numbers indicated by the tops of the bars are no longer 1000 and 1500.
4. **(A)(B)** The truncation makes the visual appearance of the difference between the data appear reversed (in the opposite direction from what it is).

Absolute vs relative numbers

5-6 (2 pts) Preamble: suppose the baseline incidence of a disease is 2 per 10,000 and that a medicine reduces it to 1 in 10,000. Which are true (rely on class discussion)? **A = True, B = False**

5. **(A)(B)** For many people, the language of the preamble makes the medicine's efficacy seem less impressive than a description using relative numbers (% cure rate).
6. **(A)(B)** For many people, the language of the preamble makes the medicine's efficacy seem less impressive than a description using absolute numbers.

7-9 (2 pts). For a headline advertising that the percent of women among all college admissions has gone down over the last decade, which of the following possibilities could be consistent with this claim? **A = Possibly consistent, B = Not**

7. **(A)(B)** The total number of women admitted has gone up.
8. **(A)(B)** The total number of women admitted has not changed.
9. **(A)(B)** The total number of women admitted has gone down.

False positives, false negatives

10-14. You will get yourself tested for covid. Assume that the test is 98% accurate – that if someone has covid, there is only a 2% chance that the test will come back negative and that if someone does not have covid, there is only a 2% chance the test will come back positive. Answer the following. **A = True, B = False**

10-12 (2.5 pts) Assume that you have covid – that you are truly positive. What is true about your expectations going into the test (before you receive any results)?

10. **(A)(B)** There is a 2% chance that your test will be negative.
11. **(A)(B)** There is a 98% chance that your test will be negative.
12. **(A)(B)** There is not enough information given to answer this question.

13-14 (1.5 pts) Now make no assumptions about whether you have covid or not – no one knows. But your test has come back positive. What is the probability that you have covid?

13. **(A)(B)** There is a 98% chance that you are truly positive – that you have covid.
14. **(A)(B)** There is not enough information given to answer this question.

Evaluation

15. (1.5 pt) A new study, just published, is the first to report occasional but statistically significant health problems caused by long term use of a dietary supplement called Factor-F. Factor-F was legally marketed without drug trials because, as a dietary supplement, trials and FDA approval was not required. This new study is in fact the first to investigate the health effects of Factor-F. Which one of the following possible consequences of this study do we expect? One answer only.

(A) The FDA (or some other government agency) will impose an immediate halt to further sales of Factor-F.

(B) Further studies will be undertaken to see if the results are repeatable before any action is taken to halt sales of Factor F.

16-19 (3 pts) What is the meaning when a study reports a 'significant' result at $P = 0.03$? (A) = true (B) = false

16 (A)(B) There is only a 3% chance that similar results would be found if the study was repeated.

17 (A)(B) The observed data are expected only 3% of the time if the null model/hypothesis is true.

18 (A)(B) The data are expected 97% of the time if the null model/hypothesis is true.

19 (A)(B) The study results should be ignored because there is only a 3% chance of them being true – they need to exceed a 5% threshold to be considered true.

Absence of evidence, evidence of absence

20-24 (3 pts) Which statements can legitimately be made in the absence of evidence (i.e., they can be made if you have no evidence)?

20. (A)(B) Bigfoot does not exist (A) = can be made in absence of evidence (B) = cannot

21. (A)(B) We do not know if Bigfoot exists

22. (A)(B) There is no evidence that GMO foods are harmful.

23. (A)(B) There is no evidence that GMO foods are NOT harmful.

24. (A)(B) The infection of nurses with Ebola virus was caused by a breach in protocol.

Correlation

25-30. (6 pts) Which of the following options is indicated in 21-26? Base your answer only on the information provided. Some questions may require you to convert the data – to look closely at what the numbers mean.

(A) no correlation or causation is described.

(B) correlation only – the statement(s) merely describe one or more non-zero correlations,

(C) causation only – the statement(s) merely describes one or more causal models

(D) correlation and causation are described and go in the same direction

(E) correlation and causation are described but go in opposite directions (Simpson's paradox)

25 (A)(B)(C)(D)(E) Application of Dr. Bell's foot lotion speeds recovery of rashes from poison ivy. People with poison ivy rashes who have used Dr. Bells foot lotion have recovered from poison ivy faster than those who did not use the lotion.

26. (A)(B)(C)(D)(E) University of Crosbyton admits women at a higher rate than men in all of its programs. Overall, more men than women are admitted to University of Crosbyton.

27 (A)(B)(C)(D)(E) 45% of UI students have student loans; the others do not have student loans. 55% of BSU students do not have student loans, the others do.

28. (A)(B)(C)(D)(E) 35% of UI students attend their university's football games

29. (A)(B)(C)(D)(E) Winter temperatures are cooler than summer temperatures.

30. (A)(B)(C)(D)(E) Studying improves exam scores. Frequent exercising also improves exam scores.

31-34 (4 pts). Answer the following questions by using this 2x2 table about which farmers have used the herbicide called 'Roundup' and whether they have a weed problem. The entries in the table are the numbers of farmers in the category:

		Does the farmer have a weed problem?	
		Yes	No
Has the farmer used Roundup?	Yes	80	115
	No	33	55

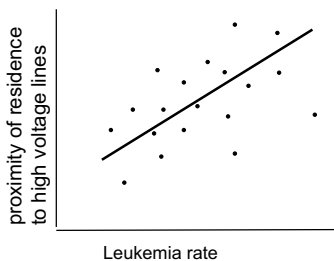
use the following answers for questions that ask for percentages:

- | | | | |
|------------------|------------------|------------------|-------------------|
| A: 0-5% | F: 25-30% | K: 50-55% | P: 75-80% |
| B: 5-10% | G: 30-35% | L: 55-60% | Q: 80-85% |
| C: 10-15% | H: 35-40% | M: 60-65% | R: 85-90% |
| D: 15-20% | I: 40-45% | N: 65-70% | S: 90-95% |
| E: 20-25% | J: 45-50% | O: 70-75% | T: 95-100% |

- 31.** For farmers that used Roundup, what is the percent with weed problems?
 (A) (B) (C) (D) (E) (F) (G) (H) (I) (J) (K) (L) (M) (N) (O) (P) (Q) (R) (S) (T)
- 32.** For farmers that did not use Roundup, what is the percent with weed problems?
 (A) (B) (C) (D) (E) (F) (G) (H) (I) (J) (K) (L) (M) (N) (O) (P) (Q) (R) (S) (T)
- 33.** What is the overall percent of farmers who have used Roundup?
 (A) (B) (C) (D) (E) (F) (G) (H) (I) (J) (K) (L) (M) (N) (O) (P) (Q) (R) (S) (T)
- 34. (A)(B)** From the table, there is a correlation between use of Roundup and having a weed problem, and the correlation is such that weed problems are more common/prevalent in farmers who have used Roundup than in those who have not used Roundup.
(A) = True (B) = false

35-37 (3 pts). Across the US we observe the pattern (shown in the figure) between proximity of a household/residence to high-voltage power lines and leukemia rate of family members in that household. Which models are *consistent* with these data? This question is the same as asking which models cannot be rejected.

A = consistent, thus cannot be rejected, B = not consistent – can be rejected



- 35.** (A) (B) Leukemia is caused by increased exposure to high voltage lines
36. (A) (B) Leukemia is suppressed by increased exposure to high voltage lines
37. (A) (B) High voltage lines have no causal effect on leukemia rates

38. (1 pt.) Exam Key Code **B:** Answer **(B)** on question **38** to indicate your exam code.
 Format: **last name first name** , last 4 digits of your V00 number , 1 answer 2 answer ...