

You must turn in this hard copy (with your name on it) and your scantron to receive credit for this exam.

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

A = True, B = False unless indicated otherwise. If any part of an answer is incorrect, treat all of it as incorrect.

Data Quality: Errors and fixes

1-4. (7 pts) Answer the questions about types of error and why.

A newspaper hires a firm to see if the responses of Republicans differ from those of Democrats to editorials published in the newspaper. The firm takes 3 surveys. Each survey uses 100 randomly chosen Democrats and 100 randomly chosen Republicans. Results of the 3 surveys are that 73%, 70% and 74% of Republicans approve of the editorials, but the corresponding numbers for Democrats are 46%, 45% and 50%. Thus there is a consistently large difference in the response between the two parties.

(A) = TRUE (B) = FALSE

1. **(A)(B)** The big difference between the parties is not likely sampling error because it is consistent across samples.
2. **(A)(B)** The big difference is not considered bias because the goal of the study is to compare the two groups; the difference is what is being measured.
3. **(A)(B)** The random choice of which Democrats to survey ensures against bias in which Democrats were chosen.
4. **(A)(B)** The difference in approval rates between Democrats and Republicans is a form of Human and Technical error.

5-7. (6 pts) Which in the following list constitute(s) bias? **(A) is bias (B) is not**

5. **(A)(B)** the display on a piece of equipment shows 2 decimal places but its sensor reads to 4 decimal places; thus the display is not as accurate as is the sensor itself.
6. **(A)(B)** a technician accidentally mixes up samples in a DNA test
7. **(A)(B)** a person unconsciously feels better than they really are because they know they took a pill that should make them well

8-10. (6 pts) Which in the following list constitute(s) human and technical (H&T) error? **(A) is H&T (B) is not**

8. **(A)(B)** A thermometer that has lost some mercury now reads 3° low
9. **(A)(B)** A piece of equipment whose display shows 2 decimal places but cannot measure to the third decimal
10. **(A)(B)** Failure of a technician to follow a protocol

11-15. (9 pts). Indicate which ideal data features are clearly described:

As an able bodied man, you are skeptical of claims that men and women are treated differently by car repair places. You decide to test whether Joe's Car Repair Shop overcharges women customers more than men, using your car. You first take the car to a diagnostics shop you trust to get a full report on engine problems and costs to repair them. Without getting any repairs done, you then take it to Joe's and ask for an estimate of engine problems and repair costs. You leave with the estimate but make no repairs. 6 months later, without having driven the car in the meantime, you have the car painted a new color, get new tires and change the car interior so that it looks different than it looked previously. You have your wife take the car to Joe's and ask for an estimate of engine problems and repair costs – without letting the mechanic know that he's seen the car previously. She brings the estimate home for comparison to the first estimate and to the trusted estimate.

(A) The feature is indicated. **(B)** not indicated or absent

11. **(A)(B)** Replication
12. **(A)(B)** Standards
13. **(A)(B)** Random
14. **(A)(B)** Blind at least one way
15. **(A)(B)** Double blind

16-18. (6 pts) Two forms of an advertisement are tested to decide which is better by giving either form to each of 200 students and then having them fill out a questionnaire that reveals how well the advertisement worked. The two forms of the ad are assigned to students randomly, and the students are unaware that there are different forms of the ad. The questionnaire is multiple choice and scored on a machine whose performance is checked initially by putting through a form whose answers are all known in advance. Overall, the responses to the two ads are slightly different, but a statistical analysis indicates that these differences are not significant. **(A) = TRUE (B) = False**

- 16. (A)(B)** The scoring of a form whose answers are known in advance constitutes a standard to check for H&T error in machine scoring
- 17. (A)(B)** That students were unaware of the different ads is sufficient to rule out all forms of bias in their responses.
- 18. (A)(B)** Sampling error in the average response to the two ads was reduced by randomly assigning the forms.

19-26. Do-it-yourself protocol. You are conducting an external review/test of a genotyping lab. Your job is to send **two tubes** to the lab, with labels. Several options are given for the content of and label on a tube. You must decide which contents to send and how to label the tubes so that the features of ideal data requested in the question are present. If a tube has a person's name on it, the lab can assume that the tube contents belong to the name of the person on the label and can infer gender but nothing else. If a tube is labeled with a number, the contents are unknown to the lab but known to you. Your options for tube contents and tube labels are:

<u>tube</u>	<u>tube label</u>	<u>Contents are from</u>	<u>Blood type</u>	<u>Gender</u>	<u>Marker type</u>
(1)	Laura Baker	Laura Baker	O	Female	+
(2)	Oz Wichman	Oz Wichman	B	Male	negative
(3)	Rachael Springman	Rachael Springman	AB	Female	+
(4)	#101	Oz Wichman	B	Male	negative
(5)	#218	Judy Timmins	O	Female	negative
(6)	#10	Pam Hines	B	Female	negative
(7)	Jerry Allison	Jerry Allison	B	Male	+
(8)	Brenda Iverson	Brenda Iverson	AB	Female	negative

In the following questions, indicate which pairs of tubes (if any) will satisfy the specified criteria.

(A) = criteria are satisfied **(B)** – not satisfied

19-22 (6pts). Which tubes achieve replication of blood type and gender but marker type is not replicated? It does not matter if the gender replication is known to the lab.

- 19. (A)(B)** tubes 2 & 6
- 20. (A)(B)** tubes 2 & 4
- 21. (A)(B)** tube 5 & at least one of the others
- 22. (A)(B)** tubes 4 & 7

23-26 (6pts). Choose two tubes so that you are certain that none of blood type, gender, marker status is replicated. Absence of replication should be blind to the lab. Assume that the lab will know gender from a name on the tube.

- 23. (A)(B)** tubes 1 & 4
- 24. (A)(B)** tubes 4 & 5
- 25. (A)(B)** tubes 5 & 7
- 26. (A)(B)** tubes 1 & 8

27-29 (6pts). The following paragraph is a description of a design. In the questions following the paragraph, mark whether the quoted text correctly indicates the data feature is present. Class size could in principle be 1 student.

An instructor ponders whether the clarity of exam text affects student performance – fuzzy text may lead to higher scores because people need to engage their brains more when reading fuzzy text. Two classes are used in the study, but they are treated differently. One class gets the test in the usual fashion, with clear, black text on a white background. The other class gets an exam that is difficult to read, with smudges that require extra scrutiny to discern the words. The instructor arbitrarily chooses the first class to get the fuzzy test. The students are informed of the study after the fact. Average scores were not different between the two classes.

A = the quote indicates the feature, B = the quote does not

- 27 (A) (B)** Replication: “two classes are used in the study, but they are treated differently”
- 28 (A) (B)** Random: “The instructor arbitrarily chooses”
- 29 (A) (B)** Blind: “an exam that is difficult to read”

30-33. (7pts) Which options identify a valid “fix” for the type of error indicated; a “fix” may either reduce that error or allow you to detect that error some of the time. **A = the fix is valid; B = the fix is not valid**

	Error	Fix
30. (A) (B)	Unintentional sample mixup during testing	Change the protocol to institute blind testing of samples
31. (A) (B)	Unintentional sample mixup during testing	Split samples when they are acquired and submit both versions of the same sample for testing
32. (A) (B)	The lab falsifies results to give the prosecution its desired results	Code the samples, leaving names off
33. (A) (B)	The lab falsifies results to give the prosecution its desired results	Randomize the samples to prevent bias

Criminal Justice

We mentioned 4 features of an ‘ideal’ forensic method for matching a suspect with a forensic sample: (i) reference database, (ii) discrete characteristics, (iii) independent verification possible, (iv) labs/experts pass blind proficiency tests (95% or better).

34-38. (8 pts) Which of the following points correctly identifies a main purpose of the feature, explains the error reduction principle or describes an indicator of the feature’s absence? **(A) = TRUE, (B) = false**

- 34. (A)(B)** Discrete characters: are needed to have a reference database
- 35. (A)(B)** Discrete characters: an indicator of its absence is the claim that a match is unique
- 36. (A)(B)** Universal protocol: helps enable other labs to independently verify results
- 37. (A)(B)** Blind proficiency tests: an indicator of its absence is the lack of a universal protocol
- 38. (A)(B)** Reference database: is needed to calculate the random match probability

39-42. (7 pts) Which of the following points are correct about methods covered in class? **(A) = TRUE, (B) = false**

- 39. (A)(B)** In eyewitness identification, it was suggested that a photo catalog of potential suspects is a valid reference database when it is used to have the eyewitness find the one picture of the individual they saw.
- 40. (A)(B)** In fingerprint matching up to the 1990s, the FBI catalogue of fingerprints was used (in most respects) as a typical reference database.
- 41. (A)(B)** FBI bullet lead analysis relied on a universal protocol used in several US labs.
- 42. (A)(B)** Most types of forensic methods used in US courts (prior to year 2000) failed on ALL four ideal features.

43-46. (7pts) Combining sources of error in forensic matches. In a court case, the forensic lab has declared a match between suspect and sample. The defense emphasizes that there are 4 completely different and unrelated possible reasons that the match may be in ‘error’ – that the suspect was not the source of the forensic sample. The defense has further come up with numbers of 0.000001, 0.002, 0.03, and 0.1 (1 in a million, 2 in a thousand, 3 in a hundred, 1 in ten) as the individual probabilities for the 4 types of error. The defense asks you for the overall chance of the match being in error – that the match is false by any combination of the 4 reasons. Which of the following are true? (A) = TRUE (B) = false

The overall chance of a ‘false’ match in this case

- 43. (A)(B)** Is the product of the 4 error probabilities
- 44. (A)(B)** Cannot be smaller than the largest of the 4 error probabilities
- 45. (A)(B)** Would change greatly if the 1 in a million error rate was reduced to 1 in a billion
- 46. (A)(B)** Cannot be determined by combining error rates from 4 *completely* different sources

47-51 (8pts). Which of the 4 features of 'ideal forensics' are indicated as being present? For all but 'independent verification', the problem must specifically describe their presence for it to be present. For 'independent verification' the problem must specifically describe it or describe a means by which independent verification could feasibly be performed by different labs. **(modified from a previous exam)**

The only evidence in a murder trial to connect the defendant to the crime is a match between the duct tape used to bind the victim and a roll of tape found in the defendant's car. The lab providing the evidence based the match on both the total number of fibers in the tape and on the fiber dimensions in the tape. The tape in the roll and on the victim both had 513 fibers in it. Also, the microscopic widths of fibers and spacings between fibers were claimed to be too similar between the victim sample and the suspect's car sample to be from different rolls. The lab analyzed 690 tape rolls from stores in 20 states in establishing the typical similarity between different rolls of tape. Although the defense attempted to find an expert to challenge this testimony, it discovered that no other lab in the world does this kind of analysis, and the lab uses a proprietary analytical method to declare a match (that it keeps secret) so that no one else can compete with its business. This court case is in fact the first time the lab has applied its method in testing whether two samples match.

A = present B= absent, incomplete or not used

- 47. (A) (B) a reference database that can be screened for a RMP is indicated
- 48. (A) (B) some of the characteristics used are discrete
- 49. (A) (B) some of the characteristics used are not discrete
- 50. (A) (B) The labs mentioned were able to pass blind proficiency tests
- 51. (A) (B) Independent verification of a declared match is possible (explicitly present or the means for doing it is described)

Data Presentation

52-56. (8 pts). Based on the Data Presentation lecture and chapter, which of the following points are true?

(A) = TRUE (B) = FALSE

- 52 (A) (B) A drug test uses a method that gives the correct answer 99% of the time. A person getting a positive result from that test could actually have a low chance that they were truly positive.
- 53 (A) (B) A drug reducing heart attacks from 4 in 100,000 patients to 1 in 100,000 patients can be advertised legitimately as causing a 75% reduction in heart attacks.
- 54 (A) (B) Data presented as natural frequencies are harder for most people to understand than data presented as conditional probabilities
- 55 (A) (B) Data presented as relative risks are often more impressive than data presented as absolute risks.
- 56 (A) (B) Graphs can change the information conveyed merely by changing the scale on the axes.

57. (4 pts.) Exam Key Code A: **Fill in** bubble (A) on question 57 to indicate your exam code; leave the other bubbles blank. Also, fill in the correct bubbles for your name and EID on the scantron form.