C - Exp Design and Data Analysis C - Nov 7th Practice - Nov 7 Country-wide SO Practice - 11-07-2020

This event will work in 2 parts.

The first part is similar to the old rules of the first part of experimental design. You will be given the materials list with images and you will only design the experiment.

At the end of the test, please remember to fill out the following form with the procedure diagrams within 15 minutes of the block ends.

PLEASE FILL THE FORM OUT AFTER THE END OF THE TEST WITHIN 15 MINUTES. IF YOU LEAVE THE BROWSER DURING THE EXAM THAT WILL BE COUNTED AS CHEATING.

https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)
OR http://tinyurl.com/so-exp-design (https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)

The second part is a test where you will be tested on data analysis as well as experimental design procedures and best practices.

Both parts will be given at the start and turned in together at the end of 50 minutes.

If you have any urgent questions during the exam you may contact the chat.

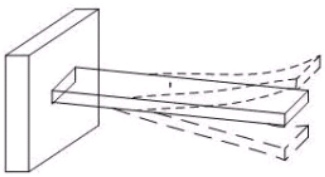
Good luck!!!

Test Writers: Peter Zhu, peterzhu@berkeley.edu

PART 1: Experimental Design

Topic Area: Beam Vibration

A straight, horizontal cantilever beam under a vertical load will deform into a curve. When this force is removed, the beam will return to its original shape; however, its inertia will keep the beam in motion. Thus, the beam will vibrate at its characteristic frequencies.



Materials:

(item 1) 1 wall with opening of size w by t where:

- w = 10 cm, t = 3 cm
- The beam can move into and out of the wall (there is adequate space for the entire beam to fit through the wall)
- The wall is stable and will not deform under any loading

(item 2) 1 steel beam of length l, width w, thickness t where:

- l = 100 cm, w = 10 cm, t = 3 cm
- The beam can move into and out of the wall (there is adequate space for the entire beam to fit through the wall)
- The beam should not fracture under any reasonable load (only elastic deformation)
- Constant material property across entire beam

(item 3) Adjustable Dumbbell Set

- Ranges from masses of 10 to 100 kg in increments of 10 kg

Measurement Equipment:

1 timer

- Measures to the nearest 0.01 seconds

1 linear measuring device

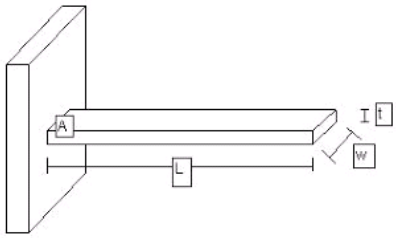
- Measures to the nearest 0.1 inches

Design an experiment that addresses the topic area of beam vibration. You must use at least 2 materials not including the measurement equipment on the list above. Your design report will include a statement of the problem, hypothesis, independent variables, dependent variables, controlled variables, constants, experimental control, materials list of the materials you used, and procedure with diagrams.

Item 3:

Item 1 (left portion - wall) and Item 2 (right portion - beam):

- Point A is the location where the beam meets the wall
- L = Length, W = Width, t = thickness



1. (2.00 pts) A. Statement of Problem

Expected Answer:

2. (6.00 pts) B. Hypothesis

Expected Answer:

3. (8.00 pts) C. Independent Variable

Expected Answer:

4. (4.00 pts) C. Dependent Variable

Expected Answer:

5. (4.00 pts) C. Controlled Variables

Expected Answer:

6. (4.00 pts) C. Constants

Expected Answer:

7. (4.00 pts) D. Experimental Control (Standard of Comparison)

Expected Answer:

8. (4.00 pts) E. Materials

Expected Answer:

9. (14.00 pts)
F. Procedure and Set-up Diagrams

At the end of the test, please remember to fill out the following form with the procedure diagrams within 15 minutes of the block ends.

PLEASE FILL THE FORM OUT ***AFTER*** THE END OF THE TEST WITHIN 15 MINUTES. IF YOU LEAVE THE BROWSER DURING THE EXAM THAT WILL BE COUNTED AS CHEATING.

https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)

OR http://tinyurl.com/so-exp-design (http://tinyurl.com/so-exp-design)

Expected Answer:

PART 2: Data Analysis

10. (1.00 pts) The standard deviation is always greater than the mean for any distribution.

○ True ◉ False

11. (1.00 pts) The variance is the standard deviation squared.

◉ True ○ False

12. (1.00 pts) The sample standard deviation is always greater than the population standard deviation.

○ True ◉ False

13. (1.00 pts) Histograms are used for categorical data.

○ True ◉ False

14. (1.00 pts) If you have an outlier, that means you made a blunder in your experiment.

○ True ◉ False

15. (1.00 pts) If the p value is less than 0.05, we must always reject the null hypothesis.

○ True ◉ False

16. (1.00 pts) A type 2 error is the rejection of a true null hypothesis.

○ True ◉ False

17. (1.00 pts) The poisson distribution is commonly used in modeling occurence of rainfall as a random event.

(●) True ( ) False

18. (1.00 pts) Brown noise is spread across the sound spectrum at all frequency levels.

( ) True (●) False

19. (1.00 pts) After removing the outliers, the max - min is the interquartile range.

( ) True (●) False

20. (1.00 pts) You are trying to measure your weight on a scale. However, the wind is blowing at you and you aren't staying very still. This is a systematic error.

( ) True (●) False

21. (1.00 pts) Accuracy is more important in experiments than precision.

( ) True (●) False

22. (1.00 pts) The variance is always greater than or equal to 0.

(●) True ( ) False

23. (1.00 pts) Suppose I have a statistical model to predict if an email is spam or not. We cannot simultaneously reduce the bias and the variance of the model.

(●) True ( ) False

24. (1.00 pts) I obtain a 90% confidence interval (30,50), but I think it is too large. I can decrease my confidence level to make the interval range smaller.

(●) True ( ) False

25. (1.00 pts) The standard error is the sample standard deviation divided by the number samples.

( ) True (●) False

26. (1.00 pts) If the correlation is -1, we expect all the residuals to be 0.

(●) True ( ) False

27. (1.00 pts) If the correlation is 0, we expect all the residuals to be 1.

( ) True (●) False

28. (1.00 pts) If all the residuals are 1, our linear regression model is incorrect.

(●) True ( ) False

29. (1.00 pts) In a right skewed distribution, the median is greater than the mean.

( ) True (●) False

30. (2.00 pts) How many sig figs does the number 94050 have?

4

31. (2.00 pts) The degrees of freedom for the Chi-Square test statistic when testing for independence in a contingency table with 10 rows and 3 columns would be

18

32. (3.00 pts) Let X be the roll of a fair die. What is the variance of X? Reduce your answer the simplest fractional form, do not include spaces in your answer.

35/12

33. (4.00 pts) Assume that a dimensionless quantity Y is expressed by two dimensionless quantities X, Z as follows.

$Y = XZ + Z$

It is known that $X = 1.0 \pm 0.1$ and $Z = 3.0 \pm 0.2$ at 90 % confidence. Based on Gauss' method of error propagation, what is the uncertainty in Y?

Answer is decimal form. Your answer should be "0.XX", meaning include 2 decimal places and the "0." in your answer.

0.50

34. (3.00 pts) Which one can be the cause of bias errors of measurements?

(Mark **ALL** correct answers)

- [ ] A) Random incorrect reading of scale
- [ ] B) Uncompensated vibration
- [x] C) Imperfect calibration

35. (2.00 pts) Given Var(X+Y) = Var(X) + Var(Y), what is the correlation of X and Y, Corr(X, Y)?

0

36. (3.00 pts) Which of the following distributions is appropriate for modeling the number of heads in 10 filps of the same coin?

(Mark **ALL** correct answers)

- [ ] A) Exponential
- [x] B) Binomial
- [ ] C) Geometric
- [ ] D) Bernoulli

- [ ] E) Chi-Squared
- [ ] F) Student's t

The following 4 questions are related to the the following scenario below.

On Yelp, you find an article claiming that 20% of San Diego restaurant menu items are appetizers, 30% are desserts, and 50% are main dishes. Your friend says, “I don’t care about desserts or main dishes, but I think I’ve seen much more than 20% of menu items that are appetizers in San Diego! I think that number is too low.” Note that every item in a menu fall under appetizers, main dishes, or desserts and cannot fall in multiple categories. You decide to collect data on a simple random sample of 100 menu items from local restaurants in San Diego and do some hypothesis testing. You obtain the following data:

Appetizers: 31, Main Dishes: 40, Desserts, 29

For this question, we will be verifying your friend's clain using the data collected.

37. (3.00 pts) Which of the following could be a valid null hypothesis for testing your friend's claim?

(Mark **ALL** correct answers)

- [ ] A) Appetizers are more popular in San Diego than other parts of the United States.
- [ ] B) The Yelp article you read is incorrect.
- [x] C) Menu items in your data are a random sample from a distribution with a 20% chance of being an appetizer, a 30% chance of being a dessert, and a 50% chance of a main dish.
- [ ] D) In a random sample of 100 menu items, there should be an equal number of appetizer and non-appetizer items. Any deviation from that is due to chance.
- [ ] E) In a random sample of 100 menu items, we should see exactly 20 appetizers, 30 desserts, and 50 main dishes.
- [x] F) Menu items in your data are a random sample from a distribution with a 20% chance of being an appetizer and a 80% chance of being a non-appetizer.

38. (3.00 pts) Which of the following could be a valid alternative hypothesis for testing your friend's claim?

(Mark **ALL** correct answers)

- [ ] A) Appetizers are more popular in San Diego than other parts of the United States.
- [x] B) The chance that a random menu item in San Diego is an appetizer is more than 20%
- [ ] C) The chance that a random menu item in San Diego is an appetizer is less than 20%
- [ ] D) The chance that a random menu item in San Diego is an appetizer is different from 20%
- [ ] E) The data collect do support your friend's claim.
- [ ] F) The distribution of menu items in San Diego is different from the distribution of menu items in the United States.

39. (3.00 pts) What is a good choice of test statistic for this test?

(Mark **ALL** correct answers)

- [x] A) In a sample of 100 menu items, the number of appetizer items
- [x] B) In a sample of 100 menu items, the number of appetizer items minus 20
- [ ] C) The absolute value of the previous option
- [ ] D) In a sample of 100 menu items, the number of restaurants with appetizer items

40. (3.00 pts)
Let's say you run a simulation and obtain 100 simulations under the null hypothesis. Which of the following statements is true about your p value from your experiments?

(Mark **ALL** correct answers)

- [x] A) If the p-value is exactly .01, then 1 of the simulations produced test statistics greater than or equal to the one observed in the data

- [ ] B) The p-value must be less than .05
- [x] C) A p-value close to 1 means that the data are consistent with the null hypothesis
- [x] D) A p-value close to 0.5 means that the data are consistent with the null hypothesis
- [ ] E) A p-value close to 0 means that the data are consistent with the null hypothesis
- [ ] F) The p-value is the probability that your friend is right.

The following 4 questions are related to the the following scenario below.

Consider the following: “You are eligible to serve as a juror if you are 18 years old, a U.S. citizen and a resident of the county or district where summoned. You must be able to understand English, and be physically and mentally capable of serving. In addition, you must not have served as any kind of juror in the past 12 months, must not be currently incarcerated in any prison or jail, and must not have been convicted of a malfeasance in office for which your civil rights have not been restored.“

Consider the following: “The names of jurors are selected at random from lists of registered voters. The law also allows that courts may use the names of all persons who have driver licenses or identification cards issued by the Department of Motor Vehicles.”

41. (2.00 pts) This method of sampling leads to non-response bias because the sampling frame is smaller than the eligible population.

- ○ True
- ◉ False

42. (2.00 pts) This method of sampling leads to selection bias because the sampling frame is different from the eligible population.

- ◉ True
- ○ False

43. (2.00 pts) This method of sampling leads to response bias because the sampling frame is bigger than the eligible population.

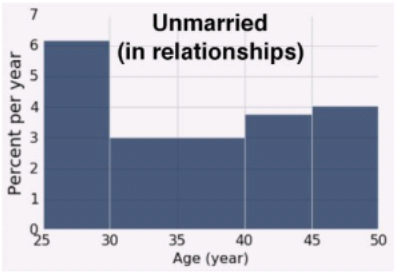
- ○ True
- ◉ False

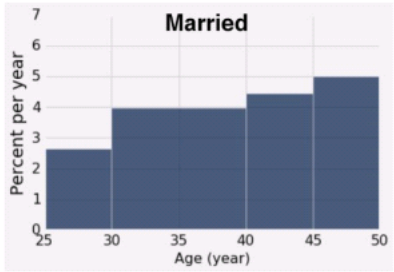
44. (3.00 pts)
The described process of random selection is carried out. The data show that the average educational level of the selected jurors is higher than the average educational level of the eligible population. This is due to

- ○ A) chance error but not bias
- ○ B) bias but not chance error
- ○ C) neither bias nor chance error
- ◉ D) both bias and chance error

The following 5 questions are related to the the following scenario below.

500 women age 25 to 49 in steady relationships were surveyed. Each woman was asked her age in years and whether she was married to her partner. There were 400 unmarried and 100 married women among those surveyed. The histograms below visualize the ages of these two groups of women.





For each pair of quantities, compare them using the information above and choose one of following:

45. (2.00 pts) (I) The number of unmarried women age 25-29 vs (II) The number of unmarried women age 30-39

- A) (I) is larger.
- B) (II) is larger
- C) (I) and (II) are about the same.
- D) There is not enough information to compare (I) and (II).

46. (2.00 pts) (I) Among the unmarried women, the proportion who are of age 25-29 vs (II) Among the married women, the proportion who are of age 45-49

- A) (I) is larger.
- B) (II) is larger.
- C) (I) and (II) are about the same.
- D) There is not enough information to compare (I) and (II).

47. (2.00 pts) (I) The number of unmarried women age 30-39 vs (II) The number of married women

- A) (I) is larger.
- B) (II) is larger.
- C) (I) and (II) are about the same.
- D) There is not enough information to compare (I) and (II).

48. (2.00 pts) (I) The proportion of married women age 30-34 vs (II) The proportion of married women age 35-39

- A) (I) is larger.
- B) (II) is larger.
- C) (I) and (II) are about the same/
- D) There is not enough information to compare (I) and (II).

49. (2.00 pts) (I) The 20th percentile age of unmarried women vs (II) The 20th percentile age of married women

- A) (I) is larger.
- B) (II) is larger.
- C) (I) and (II) are about the same.
- D) There is not enough information to compare (I) and (II).

At the end of the test, please remember to fill out the following form with the procedure diagrams within 15 minutes of the block ends.

PLEASE FILL THE FORM OUT AFTER THE END OF THE TEST WITHIN 15 MINUTES. IF YOU LEAVE THE BROWSER DURING THE EXAM THAT WILL BE COUNTED AS CHEATING.

https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)

OR http://tinyurl.com/so-exp-design (https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)
(https://docs.google.com/forms/d/e/1FAIpQLSciCwAibQL45VzjVVvhUDize2WOrJAL5Y9nQ1YO1aZE-ksggQ/viewform)