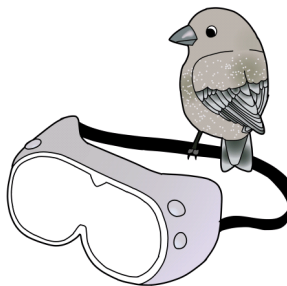**Experiment and Data Analysis C - Experimental Design C - BirdSO C 2021 Invitational - 03-07-2021**

Instructions (shown before students start the test)

Welcome to the 2021 BirdSO Experiment and Data Analysis C Exam!

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. 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.

Introduction (shown after students start the test)

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

PART 1: Experimental Design

Topic Area: Material Fatigue

Definition: Fatigue is a type of failure caused by cyclic loading at stress levels below those of failure under static loading. Fatigue damage is a cumulative process comprising of three stages: crack initiation, crack growth, and fracture.

Materials:

(item 1) 3 Identical Rubber Bands

(item 2) 1 Slotted Mass Set with Hook. Each mass is 100 g, and there are 10.

(item 3) 1 roll of tape

(item 4) 1 plastic bag

(item 5) 1 pair of scissors

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 material fatigue. 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 1 | Item 2 | Item 3 | Item 4 | Item 5 |

---

1. (2.00 pts)   A. Statement of Problem

---

2. (6.00 pts)   B. Hypothesis

---

3. (8.00 pts) C. Independent Variable

4. (4.00 pts) C. Dependent Variable

5. (4.00 pts) C. Controlled Variables

6. (4.00 pts) C. Constants

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

8. (4.00 pts) E. Materials

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

PART 2: Data Analysis

10. (2.00 pts)
Researchers are studying the effectiveness of a particular vaccine. A large random sample was taken from the population of people who took the vaccine in 2018. Among the sampled people, 76% did not get the disease. Another large random sample was taken in 2019 among the people who took the vaccine that year, and 40% did not get the disease.

A statistician thinks the vaccine was less effective in 2019 than in 2018.

Which of the following could be the correct <u>null</u> hypothesis? (There may be 1, or more, but at least 1)

(Mark **ALL** correct answers)

☐ A) The vaccine was less effective in the 2019 population than in the 2018 population, due to chance.

☐ B) The vaccine was equally effective in the two samples but its effectiveness was different in the two populations due to chance.

☐ C) The vaccine was equally effective in the two populations but its effectiveness was different in the two samples due to chance.

11. (2.00 pts)   A _____ corresponds to a decrease in the width of a confidence interval?

(Mark **ALL** correct answers)

☐ A)   increase in confidence level

☐ B)   decrease in confidence level

☐ C)   increase in sample size

☐ D)   decrease in sample size

12. (0.50 pts)   How many significant figures are in the number 100? Answer as an integer below.

13. (0.50 pts)   How many significant figures are in the number 101? Answer as an integer below.

14. (0.50 pts)   How many significant figures are in the number 1010? Answer as an integer below.

15. (0.50 pts)   How many significant figures are in the number 101.010? Answer as an integer below.

[   ]

16. (1.00 pts)   Which of the following visualizations is most used to show quantitative time series data?

○ A) Scatterplot
○ B) Box Plot
○ C) Bar Graph
○ D) Line Graph
○ E) Violin Plot
○ F) Pie Chart

17. (1.00 pts)   The variance is always greater than the standard deviation.

○ True   ○ False

18. (1.00 pts)   For a sample from a normal distribution, the mode cannot be an outlier.

○ True   ○ False

19. (1.00 pts)   Bootstrapping involved resampling _____ replacement.

○ A) with
○ B) without
○ C) either with or without

20. (1.00 pts)   A bias in measurement is a systematic error.

○ True   ○ False

21. (2.00 pts)   All surveys contain response bias.

○ True   ○ False

22. (3.00 pts)

Why is the above visualization bad, and how might we fix it (ie, what other plot should we use)?

---

23. (1.00 pts) All precise measurements must be accurate.

○ True ○ False

---

24. (1.00 pts) Measurements cannot be both precise and accurate.

○ True ○ False

The following 5 questions are related to this plot below.

---

25. (1.00 pts) Which company has the largest range for number of stock sold?

○ A) A
○ B) B
○ C) C

---

26. (1.00 pts) Which company has the largest interquartile range for number of stock sold?

○ A) A
○ B) B
○ C) C

---

27. (1.00 pts) Company C had the highest average number of stocks sold.

○ True ○ False

28. (1.00 pts) Company B's distribution of number of stocks sold has positive skew.

○ True ○ False

29. (1.00 pts) The median number of stocks sold for company B was greater than that of company A.

○ True ○ False

30. (2.00 pts) Consider the list of numbers: [1 2 3 4 5 6 7 8 9 10]

What is the sample standard deviation? Your answer should be 3 significant figures!

31. (1.00 pts) Consider the list of numbers: [1 2 3 4 5 6 7 8 9 10]

Which distribution best describes this list of numbers?

○ A) Uniform
○ B) Bernoulli
○ C) Normal
○ D) Poisson
○ E) Binomial
○ F) Geometric

32. (1.00 pts) What distribution best describes the outcome a single flip of an unfair coin where there are only 2 possible outcomes?

○ A) Uniform
○ B) Bernoulli
○ C) Normal
○ D) Poisson
○ E) Binomial
○ F) Geometric

33. (4.00 pts) A bus arrives at a bus stop according to a Poisson distribution with an average of 3 arrivals every 20 minutes.

Calculate the probability that there will not be more than one bus arrival in a 20 minute time span.

34. (3.00 pts) The bias error for a measurement is 1 inch. The random error for this measurement is 0.5 inches.

What is the total uncertainty? Round your answer to 4 significant figures in inches.

35. (3.00 pts) If the sum of the residuals for a linear regression model is 1, this means that:

(Mark **ALL** correct answers)

- ☐ A) There is positive correlation
- ☐ B) The covariance is 0
- ☐ C) The model is a perfect fit
- ☐ D) The model is incorrect
- ☐ E) There are no outliers
- ☐ F) The distribution is normal

36. (1.00 pts) Sampling from a Bernoulli distribution with p=0.5 is very precise.

○ True ○ False

37. (2.00 pts) Suppose that for some variable X, we have $\sigma_X = 5$. Let Y be defined as $Y = 10X + 20$. What is σ_Y?

Write your answer as an integer below.

38. (2.00 pts) A sample with n members has a margin of error (MOE) $z \pm k$. What is the MOE of a sample with $2n$ members?

39. (1.00 pts)
Jason wants to take a sample of 15 teams out of those who competed at BirdSO. Which of the following is closest to an example of how he could obtain a sample using systematic sampling?

○ A) He inputs each team number into a random number generator and generates 15 numbers.

○ B) He picks 5 teams from Division B and 10 from Division C randomly.

○ C) He picks the 15 teams who turn in the Experimental and Data Analysis test first.

○ D) He randomizes the order of the teams, then picks every fourth team until he has 15.

The following 3 questions relate to the graph below.

40. (1.00 pts) Which of the following values could possibly be the median of the scores?

(Mark **ALL** correct answers)

- ☐ A) 1300
- ☐ B) 2000
- ☐ C) 2400
- ☐ D) 3100
- ☐ E) 6420
- ☐ F) 10000

---

41. (1.00 pts)   In which interval is the score corresponding to the 90th percentile located?

- ○ A) $[4500, 6000]$
- ○ B) $[6000, 7500]$
- ○ C) $[7500, 9000]$
- ○ D) $[9000, 10500]$

---

42. (1.00 pts)   How many scores are at most 3000?

- ○ A) 20
- ○ B) 25
- ○ C) 30
- ○ D) 35
- ○ E) 40
- ○ F) 45

---

43. (2.00 pts)

Suppose that a simple random sample of 144 individuals is taken from a population which has mean 31 and standard deviation 9. Find the approximate probability that $\bar{x}$ is within 1.5 of the true population mean μ.

Explain your reasoning for full credit.

---

44. (4.00 pts)

As a user interface designer, you have two ideas you want to test for which is better. You have two color schemes for the product Spotify (green and black, or orange and black). Your goal may be to increase the listening times for users.

How would you go about testing this to see which color scheme is better? What kind of test(s) should you use?

How can you reduce bias and error?

---

45. (3.00 pts)

Select the TRUE statements given the following residual plot

(Mark **ALL** correct answers)

- ☐ A) The sum of the residuals is slightly greater than 0
- ☐ B) The best fit equation is an appropriate fit that accurately describes the trend
- ☐ C) The x axis of the residual plot is the same as the x axis on the original graph
- ☐ D) The residual values have half the values above 0 and the other half below 0
- ☐ E) Residual values close to 0 are close to the line of best fit
- ☐ F) None of the above

---

46. (1.00 pts) The correlation and covariance are directly proportional.

○ True ○ False

---

47. (2.00 pts) There is a 1/100 chance that gamestop stock (GME) goes to the moon. Assume every day is indepdent of other days.

On average, how many days do we expect to wait until gamestop stock (GME) hits the moon? What distribution did you use to solve this?

---

48. (2.00 pts) When conducting experiments, why is studying the types of noise (white, pink, Brownian) important?

---

49. (2.00 pts)

Assume you are doing a least squares linear regression on only 2 data points. What is the correlation coefficient and coefiicent of determination? Explain your reasoning

---

50. (1.00 pts)

100 students are supposed to use calipers to measure the length of a ballpoint pen when the pen tip is not out. However, half the students measure the length when the tip of the pen is out, and the other half measure the length when the tip is not out. You may assume the measurement device of calipers are almost completely error free.

1) What kind of distribution of pen lengths do we expect from this?

---

51. (1.00 pts) 2) Comment on the precision and accuracy of the entire group of 100 students.

(Mark **ALL** correct answers)

☐ A) precise

☐ B) not precise

☐ C) accurate

☐ D) not accurate

---

52. (1.00 pts)   3) Comment on the precision and accuracy of the half of students that measured the pen when the tip is not out

(Mark **ALL** correct answers)

☐ A) precise

☐ B) not precise

☐ C) accurate

☐ D) not accurate

---

53. (4.00 pts)

10 values in a sample have an average of 5.5 and a standard deviation of 1.3. What is the 95% confidence interval for the true mean? Show your work. Above is part of a Student's T Distribution table.

---

54. (2.00 pts)   Explain how you would linearize the following data?

---

55. (1.00 pts)   We conduct some chi-square test and notice that the chi-square value is higher than the critical value at some p value. Which of the following can we do?

(Mark **ALL** correct answers)

☐ A) Accept the null hypothesis

☐ B) Reject the null hypothesis

☐ C) Accept the alternative hypothesis

☐ D) Reject the alternative hypothesis

---

56. (3.00 pts)   I flip two fair coins. Let X be the number of heads (0, 1, or 2), and let Y be an indicator for the first coin being a head (0 or 1).

What is the cov(X, Y)? Show your work.

Use the following set of data to answer the following questions:

57. (1.00 pts)   What is the name of the way this set of data is displayed?

- A) Box and Whisker Graph
- B) Stem and Leaf Plot
- C) Data Table
- D) Cumulative Frequency Table
- E) Histogram

58. (1.00 pts)   Find the mean of the data, to the nearest hundredths.

59. (1.00 pts)   Find the median of the data, to the nearest hundredths.

60. (2.00 pts)   Compare the values of the median and the mean; which one is greater than the other? What does this tell you about the data set?

61. (1.50 pts)   How many outliers are present in the data? Enter one number only.

62. (3.00 pts)   Describe how you found these outliers, as well as calculations and values that you used to come to your conclusion.

Error Propagation Marathon!

Answer to the nearest hundredths, in the format a.bc ± d.ef.

63. (1.50 pts) Find the uncertainty of x = (3±0.5) + (6±1.2)

64. (1.50 pts) Find the uncertainty of x = (12±0.3) + (100±1.4) + (55±1.2) + (3±1.2) - (3±1.1)

65. (2.00 pts) Find the uncertainty of x = (3±0.8)(3±0.7)(3±0.5)/(5±1.1)

66. (2.00 pts) Find the uncertainty of x = (3±0.75)^(2±0.5)

67. (3.50 pts) Find the uncertainty of x = (3.2±0.5)^[(3±0.65)(2.9±0.3) - (7±0.2)]

68. (4.00 pts) Find the uncertainty of x = log((0.55±0.1) * (3.1±0.15) / (1.5±0.1) + (2±0.2))

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