Fan Bu
Duke University - Department of Statistical Science

Application Exercise 2.1: Distributions of numerical variables

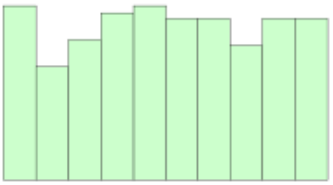
Your name: _____

Write your responses in the spaces provided below (you can also use the accompanying blank sheet if you need more space). WRITE LEGIBLY!

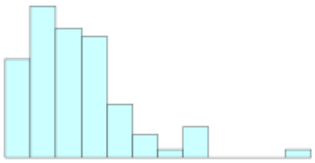
Shapes of distributions

1. Below are two histograms. One corresponds to the age at which a sample of people applied for marriage licenses; the other corresponds to the last digit of a sample of social security numbers. Which graph is which, and why?

(a)



(b)



2. Match the following variables with the histograms and bar graphs given below. These data represent Sta 101 students at Duke. [*Hint: Think about how each variable should behave.*]

   (a) the height of students

   (b) gender breakdown of students

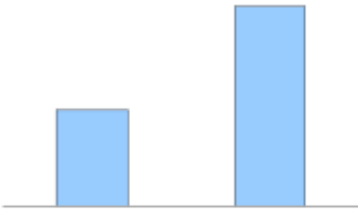
   (c) the time it took students to get to their first class of the day

   (d) the number of hours of sleep students received last night

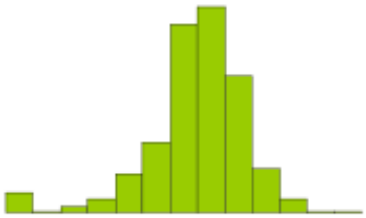
   (e) whether or not students live off campus

   (f) the number of piercings students have

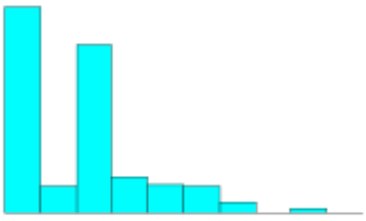
(1)



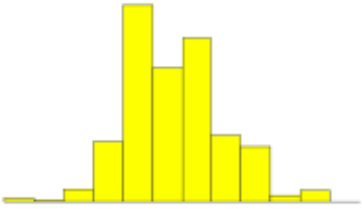
(3)



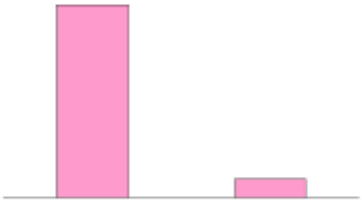
(5)



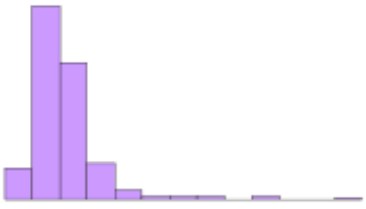
(2)



(4)



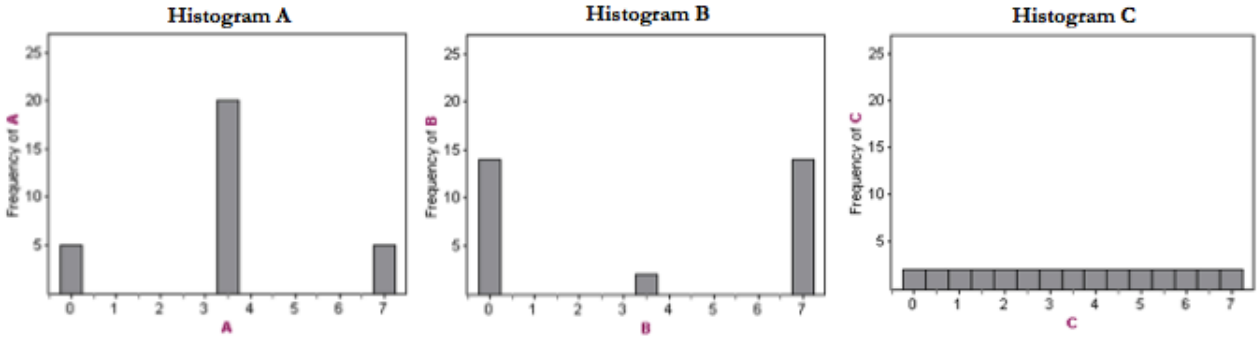
(6)



3. Come up with a concise way (1-2 sentences) to teach someone how to determine the expected distribution of any variable.

Variability

4. Order histograms A, B, and C from least to most variable. Explain your reasoning. (Hint: Variability is the extent to which data points in a statistical distribution diverge from the average as well as the extent to which these data points differ from each other. So, you'll want to think about spread and symmetry).



5. Between histograms D and E, which exhibits more variability? Explain your reasoning.