***Choose the one best answer for each question below by writing the corresponding letter in the blank to the left of the question. Each question is worth 1 point.***

**1.** What symbol is used to represent the population standard deviation?

\_\_\_\_\_ **A.**  **B.**  **C.** Q3 **D.** x **E.** s

**2.** What symbol is used to represent the population mean?

\_\_\_\_\_ **A.**  **B.**  **C.** Q3 **D.** x **E.** s

**3.** What is the name of the variable that we are interested in predicting or explaining?

\_\_\_\_\_ **A.** Continuous **B.** Discrete **C.** Explanatory **D.** Response **E.** Sampling

**4.** What is the vertical difference between an observed and predicted value of the response variable?

\_\_\_\_\_ **A.** Explanatory **B.** Frequency **C.** Predicted **D.** Residual **E.** Response

**5.** Which of the following is not described in a bivariate EDA for quantitative variables?

\_\_\_\_\_ **A.** Association **B.** Form **C.** Outliers **D.** Shape **E.** Strength

***Answer questions 6-12 in the provided space.***

1. **[4 pts]** What are the two major goals of linear regression?
2. **[3 pts]** Define RSS and explain how it is used to find the “best-fit line.”
3. **[10 pts]** Krochmal and Sparks (2007; J. Mammalogy) examined the relationship between forearm length and age of 38 Northern Myotis Bats (*Myotis septentrionalis*) captured in Vermillion County, IN. Construct a complete bivariate EDA from the results in Figure 1.

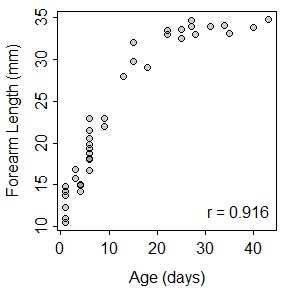


Figure 1. Scatterplot for Northern Myotis Bat forearm length on age.

1. **[10 pts]** Poysa (2003; Behav. Ecol. Sociobio.) examined the effect of predation risk on the number of eggs laid by Goldeneye ducks (*Bucephala clangula*). Specifically, she identified groups of nests that were near each other and then measured relative predation risk (a metric that ranges from 0 to 100, with 100 representing the highest predation risk) and mean number of eggs laid for each group of nests. Construct a complete bivariate EDA from her results in Figure 2.

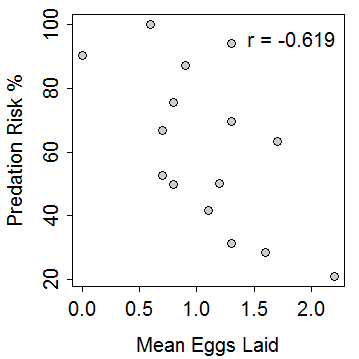


Figure 2. Scatterplot of predation risk on mean number of eggs laid for Goldeneye ducks.

1. **[16 pts]** Researchers observed groups of dolphins off the coast of Iceland in 1998. The researchers recorded time of day (Morning, Noon, Afternoon, and Evening) and the main activity of the group (travelling, feeding, or socializing). The number of dolphin groups observed in each time of day and activity is shown in Table 1. Use these results to answer the questions below the table *to one decimal place and show your work*.

Table 1. Frequency of dolphin groups by time of day and type of activity.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Type of Activity** | | |
| **Time of Day** | **Traveling** | **Feeding** | **Socializing** |
| **Morning** | 6 | 28 | 38 |
| **Noon** | 6 | 4 | 5 |
| **Afternoon** | 14 | 0 | 9 |
| **Evening** | 13 | 56 | 10 |

1. What percentage of dolphin groups observed in the evening were feeding?
2. What percentage of socializing dolphin groups were observed in the morning?
3. What percentage of all dolphin groups were observed feeding in the morning?
4. What percentage of all dolphin groups were observed in the evening?
5. **[22 pts]** A company produces packages of grape tomatoes that they advertise as weighing 16 ounces. If a package weighs less than 15.8 ounces or more than 16.5 ounces it is rejected for sale. Suppose that it is known that the distribution of package weights prior to possible rejection is normal with a mean of 16.1 ounces and a standard deviation of 0.3 ounces. Use this information to answer the questions below. *Please write the R code that you used along with your final numerical answer to one decimal place.*
6. What percentage of packages are rejected for being too heavy?
7. What percentage of packages are acceptable for sale?
8. What should the weight limit be changed to so that only 5% of packages would be rejected for being too heavy?
9. What is the IQR for weight of packages?
10. What is the median weight of packages?

**library(NCStats)**

**distrib(val,mean=meanval,sd=sdval,lower.tail=FALSE,type=”q”)**

where

* **val** is a value of the quantitative variable (x) or an area (i.e., percentage, but entered as a proportion)
* **meanval** is the population mean ()
* **sdval** is the standard deviation () or error (SE)
* **lower.tail=FALSE** is included for “right-of” calculations
* **type=”q”** is included for reverse calculations

1. **[30 pts]** Male Magnificent Frigatebirds (*Fregata magnificens*) display an enlarged red throat pouch to attract females. However, females may not always be able to see the size of the pouch. Madsen *et al.* (2004) hypothesized that females may judge the size of the pouch from the frequency of the drumming sound produced by the pouch. To determine if pouch size could be predicted from drumming frequency, Madsen *et al.* recorded pouch size and frequency of the drumming sound for 18 males (Figure 3). Use this information to answer the questions below. *Show your work as necessary*.

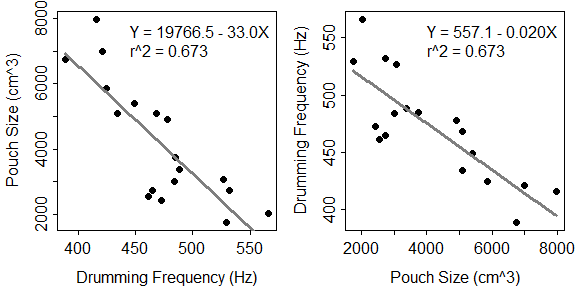


Figure 3. Fitted line plot for pouch size on drumming frequency (Left) and drumming frequency on pouch size (Right).

1. In terms of the variables of this problem, interpret the value of the slope? [*use a complete sentence*]
2. In terms of the variables of this problem, interpret the value of the intercept? [*use a complete sentence*]
3. How much would one expect pouch size to change if drumming frequency increased by 100 Hz?
4. What is the predicted pouch size if drumming frequency is 200 Hz?
5. What is the predicted pouch size if drumming frequency is 500 Hz?
6. What is the residual if pouch size is 4000 cm3 and drumming frequency is 450 Hz?
7. What percentage of variability in pouch size is explained by drumming frequency?
8. What is the correlation coefficient between pouch size and drumming frequency?
9. Do you have concerns about this regression? [*thoroughly explain your answer, whether you have any concerns or not.*]