

University of Pennsylvania
BIOL4536 Fall 2023

HW#2
(Regression)

Assigned September 6th
Due September 13th, 3:30pm

Problem 1. (3 points)

Using the formula of expected value for a discrete random variable (on slide #11 of the Regression lecture) prove that if all values in the range of a random variable X are equally likely, then $E[X]$ equals the average of the values in the range of X .

Problem 2. (4 points)

Consider the regression model

$$Y = \beta_0 + \beta_1 X + \epsilon$$

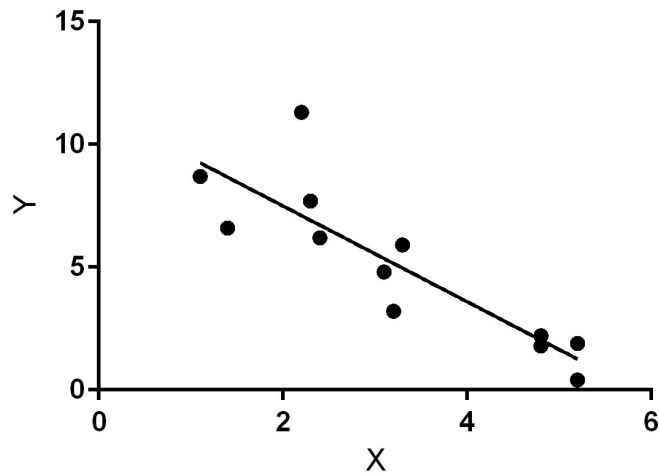
- (i) What component of this model is the source of all randomness?
- (ii) How many unknown parameters must be specified to completely specify the model?
- (iii) True or False: If $\sigma = 0$ then Y is a function of X ?
- (iv) What is the form (shape) of the regression curve of the following model?

$$Y = \beta_0 + \beta_1 2^X + \epsilon$$

Problem 3. (3 points)

Consider this data, two measurements from 12 subjects:

	X	Y
subject #1	1.4	6.6
subject #2	1.1	8.7
subject #3	2.2	11.3
subject #4	2.3	7.7
subject #5	2.4	6.2
subject #6	3.2	3.2
subject #7	3.1	4.8
subject #8	3.3	5.9
subject #9	4.8	2.2
subject #10	4.8	1.8
subject #11	5.2	1.9
subject #12	5.2	0.4



Use this Linear Regression Calculator:

<https://www.graphpad.com/quickcalcs/linear1/>

- (i) What is the regression equation?
- (ii) According to our notation, what are the estimates of β_0 and β_1 ?
- (iii) Is there a significant regression relation between the variables?
- (iv) Using the model, what is the estimate of Y when $X = 4$?
- (v) Which is definitely true:
The probability that the Y -Intercept is > 14 is
(A) ≤ 0.05
(B) ≥ 0.05