Chapter 5: Probability

Investigators recently reported the results of a study designed to assess whether two different medical treatments are effective in treating ICU patients having the same virus. The investigators were interested in whether a new medicine, Treatment 2, performed better than the placebo at getting patients to leave the intensive care unit (ICU). The investigators included Treatment 1, an older medication used to treat the virus, as another way to measure the effectiveness of the new medicine (Treatment 2). Their results are presented in the table below.

	Treatment 1	Treatment 2	Placebo	TOTALS
Left ICU after 2 weeks	35	44	26	105
Left ICU after 3 weeks	75	125	20	220
Left ICU after 4 weeks	45	65	15	125
TOTALS	155	234	61	450

1. What is the probability that a randomly selected subject left the ICU after 3 weeks?

<div align="right">1. _____</div>

2. What is the probability that a randomly selected subject was treated with Treatment 2 and left the ICU after 3 weeks?

<div align="right">2. _____</div>

3. What is the probability that a randomly selected subject left the ICU after 2 weeks given that they were treated with a placebo?

<div align="right">3. _____</div>

4. What is the probability that a randomly selected subject left the ICU after 2 weeks or 3 weeks?

<div align="right">4. _____</div>

5. What is the probability that a randomly selected subject left the ICU after 2 weeks or was treated with treatment 2?

5. _____

6. What is the probability that a randomly selected subject that was given Treatment 2 left the ICU after 2 weeks?

6. _____

7. What is the probability that a randomly selected subject that was given Treatment 1 left the ICU after 3 or 4 weeks?

7. _____

8. What is the probability that a randomly selected subject didn't leave the ICU after 2 weeks, given that the subject was treated with Treatment 1?

8. _____

Chapter 4: Correlation and Regression

9. Identify which variable is likely to represent the independent (x) variable and which represents the dependent (y) variable.

   Variable 1: The number of flu cases diagnosed in a single year

   Variable 2: the year

10. Identify which variable is likely to represent the predictor variable and which represents the response variable.

    Variable 1: The number of flu cases diagnosed in a single year

    Variable 2: the year

For each of the following pairs of variables, indicate whether you would expect a positive correlation, a negative correlation, or a correlation close to 0. Explain your choice.

11.   x = the number of golfers at the course

         y = the number of available golf carts

12. Interest rate and amount of money loaned

13. Height and weight

14. Minimum daily temperature and heating cost

Use the following information to answer the next five exercises. A vacation resort rents bikes. The resort charges an up-front fee of \$10 and another fee of \$5 an hour.

15. What are the dependent and independent variables?

16. Find the equation that expresses the total fee in terms of the number of hours the equipment is rented.

17. Graph the equation

18. What is the y-intercept? Interpret its meaning using complete sentences.

19. What is the slope? Interpret its meaning using complete sentences.

Use the following information to answer the next three questions. Due to erosion, a beach shoreline is losing several thousand pounds of soil each year. A linear equation that expresses the total amount of soil lost per year is $y = 14,000x$.

20. What are the dependent and independent variables?

21. How many pounds of soil does the shoreline lose in a year?

22. What is the y-intercept? Interpret its meaning using complete sentences.

Use the following information to answer the next two exercises. The price of a single issue of stock can fluctuate throughout the day. A linear equation that represents the price of stock for Umbrella Pharmaceutical is $y = 15 – 1.5x$ where x is the number of hours passed in an eight-hour day of trading.

23. What is the slope? Interpret its meaning using complete sentences.

24. What is the y-intercept? Interpret its meaning using complete sentences.

Use the following information to answer the next two exercises. An electronics retailer used regression to find a simple model to predict sales growth in the first quarter of the new year (January through March). The model is good for 90 days, where x is the day. The model can be written as follows: $\hat{y} = 101.32 + 2.48x$ where $\hat{y}$ is in thousands of dollars.

25. What would you predict the sales to be on day 60?

26. What would you predict the sales to be on day 90?

27. Use the information in the table below to answer parts a through m. We are interested in whether there is a relationship between the <u>ranking</u> of a state and the <u>area</u> of the state.

State	Number of letters in name	Year entered the Union	Ranks for entering the Union	Area (square miles)
Alabama	7	1819	22	52,423
Colorado	8	1876	38	104,100
Hawaii	6	1959	50	10,932
Iowa	4	1846	29	56,276
Maryland	8	1788	7	12,407
Missouri	8	1821	24	69,709

(table continues)

State	Number of letters in name	Year entered the Union	Ranks for entering the Union	Area (square miles)
New Jersey	9	1787	3	8,722
Ohio	4	1803	17	44,828
South Carolina	13	1788	8	32,008
Utah	4	1896	45	84,904
Wisconsin	9	1848	30	65,499

(a) What are the independent and dependent variables?

(b) Make a scatter plot of the data.

(c) Does it appear from inspection that there is a relationship between the variables? Why or why not?

(d) Calculate the least-squares line. Put the equation in the form of: $\hat{y} = a + bx$

(e) Find the correlation coefficient. Describe the type of correlation

(f) Find the estimated areas for Alabama and for Colorado. Are they close to the actual areas?

(g) Does it appear that a line is the best way to fit the data? Why or why not?

(h) Are there any outliers?

(i) Use the least squares line to estimate the area of a new state that enters the Union. Can the least-squares line be used to predict it? Why or why not?

(j) Delete "Hawaii" and substitute "Alaska" for it. Alaska is the forty-ninth, state with an area of 656,424 square miles.

(k) Calculate the new least-squares line.

(l) Find the estimated area for Alabama. Is it closer to the actual area with this new least-squares line or with the previous one that included Hawaii? Why do you think that's the case?

(m) Do you think that, in general, newer states are larger than the original states?

Suppose that the probability distribution for the number of years it takes to earn a Bachelor of Science (B.S.) degree is given as in the table below. Use the distribution table to answer the following questions.

x	3	4	5	6	7
P(x)	0.05	0.40	0.30	0.15	0.10

28. What does x represent?

29. On average, how many years do you expect it to take for an individual to earn a B.S.?

30. Find P(5)

31. Find $P(x \leq 5)$

32. Find $P(x > 3)$

33. Find $P(x = 2)$

34. Find the probability that a student takes **at least** five years to earn a Bachelor of Science (B.S.) degree.

35. Find the probability that a student takes **not more than** four years to earn a Bachelor of Science (B.S.) degree.

36. Find the probability that a student takes **at most** five years to earn a Bachelor of Science (B.S.) degree.

37. Find the probability that a student takes **more than** four years to earn a Bachelor of Science (B.S.) degree.

38. Find the probability that a student takes **no less than** five years to earn a Bachelor of Science (B.S.) degree.

39. Find the probability that a student **does not exceed** six years to earn a Bachelor of Science (B.S.) degree.

40. Find the probability that a student earns a Bachelor of Science (B.S.) degree in **under** four years.

41. Find the probability that a student takes **over** 4 years to earn a Bachelor of Science (B.S.) degree.

42. Find the probability that a student takes **fewer than** 5 years to earn a Bachelor of Science (B.S.) degree.

43. What are the mean and standard deviation for the distribution?