# Chapter 4 Homework: Correlation and Prediction

Answer questions 1 to 3 from the following information:

> *Professor Jones has created a formula with which students can predict their grade on the final examination based on their total scores on the first two examinations. The prediction equation is Y' = 2X + 30, where X is the total number of points the student has earned on the first two tests. The final examination has these characteristics: mean = 300 and standard deviation = 100. The correlation between total score on the first two tests and the final is .80. Kelly scored a total of 60 on the first two tests.*

1. Predict Kelly's score on the final examination.

2. What is the SEE for predicting the final examination from the first two tests?

3. How likely is it that Kelly will score greater than or equal to 114 on the final examination if Kelly had a total of 60 points on the first two tests?

4. Assume that a test had the following characteristics: mean of $X$ = 50; mean of $Y$ = 500; standard deviation of $X$ = 10; standard deviation of $Y$ = 100; $r$ = .80. Calculate the standard error of estimate.

5. The line of best fit for predicting $Y$ from $X$ is defined as $Y' = 10*X + 40$. Jamie scored 40 on $X$. What is Jamie's predicted score on $Y$?

6. Assume that the SEE is 60 and that Jamie is predicted to score 440. How likely is it that Jamie's actual score on $Y$ will be greater than 500?

Answer items 7 to 9 based on the following information:
The mean for $VO_2$max is 50 ml/kg/min and the standard deviation is 10 ml/kg/min. Professor Mill desires to estimate $VO_2$max from a field test of running a given distance for time. The coefficient of determination for the correlation between $VO_2$max and running time is .84. Assume that his prediction equation for predicting $VO_2$max from running time is $VO_2$max' = time in minutes × (–10) + 125.

7. Calculate Terry's estimated $VO_2$max' if Terry's run time is exactly 10 minutes.

8. Calculate the error for predicting $VO_2$max from running time.

9. How likely is it that Terry will have an actual $VO_2$max less than or equal to 29 if Terry's predicted running time is exactly 25 minutes and the SEE is 4 minutes? The answer is a probability (i.e., a number).

Answer items 10 to 12 from the following information:

*The mean for skinfold fat (hydrostatically determined) is 25% and the standard deviation is 10%. The coefficient of determination for skinfold fat and hydrostatically determined body fat is .91. Assume that the prediction equation for predicting hydrostatically determined percent body fat from skinfolds is percent body fat' = skinfolds × 0.5 + 2.*

10. Calculate Terry's estimated hydrostatically determined percent fat if Terry's skinfold thickness is 20 mm.

11. Calculate the error for predicting percent body fat from skinfolds.

12. How likely is it that Terry will have an actual percent body fat greater than 14 if Terry's skinfold thickness is 20 mm? The answer is a probability (i.e., a number).

Answer items 13 to 15 from the following information:

*Professor Dunn is interested in predicting class success on the final examination from scores on the midterm examination. He notes that the correlation between the midterm and final is .80 (a high correlation). Assume that the tests had the following characteristics: mean of midterm = 100; mean of final = 200; s of midterm = 10; s of final = 100; r = .80.*

13. Calculate the standard error of estimate.

14. The equation for predicting final examination scores from midterm scores is defined as $Y' = 3 \times X + 10$. Terry scored 40 on the midterm. What is Terry's predicted score on the final?

15. Using the same prediction equation how likely is it that Terry's actual score on the final examination will be greater than 70? Your answer should be a number.

Use the following information to answer items 16 to 18:

*Professor Dunn is interested in predicting class success on the final exam from scores on the midterm. He notes that the correlation between the midterm and final is .60 (a moderate correlation). Assume that the tests had the following characteristics: mean of midterm = 90, s of midterm = 5, mean of final = 150, and s of final = 100.*

16. Calculate the standard error of estimate.

17. The equation for predicting final exam scores from midterm scores is defined as $Y' = 2 \times X + 40$. Kelly scored 50 on the midterm. What is Kelly's predicted score on the final?

18. How likely is it that Kelly's actual score on the final will be greater than 220?

Use the following information to answer items 19 to 21:

*Professor Kelly wants to develop a prediction equation to estimate final examination scores from a variety of class assignments (e.g., homework problems, study time, quizzes, midterm examination, class participation, book reports, number of absences). The correlations that Professor Kelly obtained are listed here:*
*a) Homework problems, r = –.30*
*b) Study time, r = .60*
*c) Quizzes, r = .58*
*d) Midterm examination, r = .72*
*e) Class participation, r = 1.10*
*f) Book reports, r = .10*
*g) Class absences, r = –.78*

19. Which is the one Professor Kelly should use and why?

20. The prediction equation that Professor Kelly developed for the midterm examination was $Y' = 3 \times X + 20$. Terry scored 30 on the midterm examination. What is Terry's predicted score on the final examination based on the equation and Terry's midterm examination score?

21. Place the correlations Professor Kelly has calculated in order from highest correlation to lowest.

22, Differentiate between a positive and negative correlation. What is the difference? Give an example of each.

23. What is the major difference between simple linear correlation and multiple regression (multiple correlation)?

Answer items 24 and 25 from this information:

*Predicted $VO_2 = 30 + 1 \times (X)$, where X is the time in minutes to complete a 1.5-mile run and the SEE is 5 ml/kg/min.*

24. If John runs the 1.5-mile run in 10 minutes, what is his predicted $VO_2$?

25. How likely (provide a probability in number form) is it that John's actual $VO_2$ is less than 35?

Answer items 26 to 28 from the following information:

*Predicted $VO_2 = 40 + 1 \times (X)$, where X is the time in minutes to complete a 1.5-mile run and the SEE is 5 ml/kg/min.*

26. If John runs the 1.5-mile run in 10 minutes, what is his predicted $VO_2$?

27. How likely (provide a probability in number form) is it that John's actual $VO_2$ is greater than 60?

28. What percentage of people will have actual $VO_2$ values between 45 and 55 if they run the 1.5-mile run in 10 minutes?

Answer items 29 and 32 from the following information:

> *The Dallas Morning News of February 22, 1999, reported the following: Question: "Will I grow up to be short like my mother or tall like my father?" Answer: "Your adult height will probably be somewhere in between. . . . Another rule of thumb uses your parents' heights. Add them, divide by two, then add three inches for a boy or subtract three inches for a girl. The result is said to be correct within two inches about 95% of the time."*

29. Create a regression equation to predict one's adult height from mother's height in inches (M), father's height in inches (F), and gender (G).

30. If the regression equation to predict one's adult height from mother's height in inches (M), father's height in inches (F), and gender (G) is (M + F)/2 + 3G, where G = 1 for a boy and G = –1 for a girl, what is the standard error of estimate (SEE)?

31. Assume that your father is 80 inches tall and your mother is 70 inches tall. What is your predicted adult height if you are (c) a boy, and (d) a girl?

32. How likely is it that your adult height will actually be less than 76 inches if you are a boy and less than 70 inches if you are a girl?

Answer items 33 to 35 from the following correlation matrix.

> *M1 through M7 indicate 7 different measures that were taken.*

| | **WinPct** | **M1** | **M2** | **M3** | **M4** | **M5** | **M6** | **M7** |
|---|---|---|---|---|---|---|---|---|
| WinPct | 1.00 | | | | | | | |
| M1 | 0.935 | 1.00 | | | | | | |
| M2 | –0.829 | –0.758 | 1.00 | | | | | |
| M3 | 0.685 | 0.683 | –0.536 | 1.00 | | | | |
| M4 | –0.129 | –0.051 | 0.182 | 0.356 | 1.00 | | | |
| M5 | 0.537 | 0.372 | –0.561 | 0.265 | 0.071 | 1.00 | | |
| M6 | –0.069 | 0.055 | 0.167 | –0.645 | –0.367 | 0.461 | 1.00 | |
| M7 | –0.259 | –0.343 | 0.610 | –0.489 | –0.252 | –0.184 | –0.259 | 1.00 |

33. What is the best single predictor of WinPct?

34. Which two variables have the lowest correlation?

35. Which variable accounts for roughly 29% of the variance of WinPct?

36. What are the two attributes of a correlation coefficient?

37. The PPM assumes the variables are _____ related.

38. What happens to the SEE as the correlation between the criterion and the predictors goes up?