

Chapter 9

Adjusting for a Factor

9.1 Study Suggestions

Simpson's paradox and standardized rates constitute a nice completion to the argument, begun in Chapter 7, that it is difficult to interpret the findings of an observational study.

The real importance of Simpson's paradox rests in the following realization. The 2×2 table of any observational study can be viewed as a 'collapsed' table. The researcher must always remember that there could exist a meaningful factor for which 'uncollapsing' the table would result in a reversal of the direction of the relationship between the response and populations.

The main problem my students have with Simpson's paradox and standardized rates is in keeping the roles of the three variables straight. There is a response, there is a variable, that defines the populations to be compared, and there is a factor that may influence the response. For example, for the data on page 295 of the text, the response is the attitude toward drinking and driving. The populations consist of female and male drivers. Finally, it was conjectured that drinking frequency would influence response, so the factor was taken to be drinking frequency.

Simpson's paradox can occur, and standardized rates will yield interesting insight, only if the factor is strongly related (statistically) to both the population and the response. For example, if women and men had the same distribution of drinking frequencies then computing the standardized rates would be a waste of time. (If men and women have the same weights, the standardized rate would equal the original rate, since there would be no effect of replacing

the women's weights with the men's weights, or vice versa, since the weights are identical.) Moreover, if attitude were not related to drinking frequency, then the proportion of successes, for either gender, would be the same in each drinking frequency category. Thus, the weights used would have no impact, since we would be computing weighted averages of identical numbers.

If you use the Mantel-Haenszel test to analyze data from an observational study, remember that there is not necessarily a single correct analysis strategy. Consider, for example, the study of how attitude towards drinking and driving depends on gender that is analyzed in the text in Chapter 9. If a researcher simply wants to compare females and males there is no need to consider using the techniques discussed in Chapter 9. If, however, the researcher wants to investigate the extent to which differences in attitude are due to differences in the frequency of drinking alcohol, then the techniques of Chapter 9 may prove helpful. In this latter case, the researcher will need to exercise judgement to decide if subgroups should be analyzed separately, or if evidence from subgroups can be combined via the MH test.

9.2 Solutions to Odd-Numbered Exercises

5. The standardized value of the test statistic for Fisher's test is

$$z = \frac{\sqrt{2,328}[634(410) - 498(787)]}{\sqrt{1,132(1,197)(1,419)(908)}} = -4.82.$$

The approximate P-value for the third alternative is twice the area to the right of 4.82 under

the standard normal curve. This area is smaller than 0.0004.

9. Summing the entries in the a , $E(A)$, and $\text{Var}(A)$ columns yields $U = 634$, $E(U) = 679.73$, and $\text{Var}(U) = 130.91$, respectively. Thus, the standardized value of the test statistic for the MH test is

$$z = \frac{634 - 679.73}{\sqrt{130.91}} = -4.00.$$

The approximate P-value for the third alternative is twice the area to the right of 4.00 under the standard normal curve. This area is smaller than 0.0004.

9.3 Exam Questions

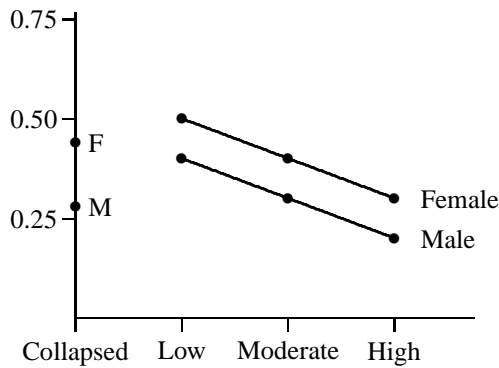
Questions 1–12 refer to the following study.

A researcher has independent random samples of 100 females and 100 males. Each subject yields a response of success or failure. The collapsed table of 200 individuals is divided into three component tables on the basis of a factor—frequency of drinking alcoholic beverages—that has possible values low, moderate, and high. Preliminary analysis yields the following table of weights and sample proportions of successes.

Factor Value	Females	
	\hat{p}_1	w_1
Low	$\hat{p}_{1L} = 0.5$	$w_{1L} = 0.6$
Moderate	$\hat{p}_{1M} = 0.4$	$w_{1M} = 0.2$
High	$\hat{p}_{1H} = 0.3$	$w_{1H} = 0.2$

Factor Value	Males	
	\hat{p}_2	w_2
Low	$\hat{p}_{2L} = 0.4$	$w_{2L} = 0.2$
Moderate	$\hat{p}_{2M} = 0.3$	$w_{2M} = 0.4$
High	$\hat{p}_{2H} = 0.2$	$w_{2H} = 0.4$

- How many females in the sample have a low drinking frequency?
- It can be shown that 44 females and 28 males gave success for their response. Use this information to draw an interaction graph of the proportion of successes. Remember to include the collapsed table in your graph.
- Which of the following expressions represents the proportion of successes in the sample of 100 females?
 - $0.5(0.6) + 0.4(0.2) + 0.3(0.2)$
 - $0.5(0.2) + 0.4(0.4) + 0.3(0.4)$
 - $0.4(0.6) + 0.3(0.2) + 0.2(0.2)$
 - $0.4(0.2) + 0.3(0.4) + 0.2(0.4)$
 - None of the above
- Which of the following expressions estimates the proportion of successes females would yield if the females in the sample had the same drinking frequency as the males in the sample?
 - $0.5(0.6) + 0.4(0.2) + 0.3(0.2)$
 - $0.5(0.2) + 0.4(0.4) + 0.3(0.4)$
 - $0.4(0.6) + 0.3(0.2) + 0.2(0.2)$
 - $0.4(0.2) + 0.3(0.4) + 0.2(0.4)$
 - None of the above
- Which of the following expressions represents the proportion of successes in the sample of 100 males?
 - $0.5(0.6) + 0.4(0.2) + 0.3(0.2)$
 - $0.5(0.2) + 0.4(0.4) + 0.3(0.4)$
 - $0.4(0.6) + 0.3(0.2) + 0.2(0.2)$
 - $0.4(0.2) + 0.3(0.4) + 0.2(0.4)$
 - None of the above
- Which of the following expressions estimates the proportion of successes males would yield if the males in the sample had the same drinking frequency as the females in the sample?
 - $0.5(0.6) + 0.4(0.2) + 0.3(0.2)$
 - $0.5(0.2) + 0.4(0.4) + 0.3(0.4)$
 - $0.4(0.6) + 0.3(0.2) + 0.2(0.2)$
 - $0.4(0.2) + 0.3(0.4) + 0.2(0.4)$
 - None of the above
- Below is the interaction graph for these data.



True or false? These data provide an example of Simpson's paradox.

8. Refer to the interaction graph for these data that is given in question 7.

True or false? The differences between female and male success proportions is the same for every level of the factor.

9. Refer to the interaction graph for these data that is given in question 7.

True or false? Compared to the component tables, the collapsed table exaggerates or amplifies the difference between females and males.

10. A researcher decides to perform an MH test on these data and obtains the following results.

Block	a	$E(A)$	$Var(A)$
Low	30	28.50	3.788
Moderate	8	6.67	3.013
High	6	4.67	2.426

Use the standard normal curve to obtain an approximate P-value for Fisher's test and the first alternative for the data from the low block.

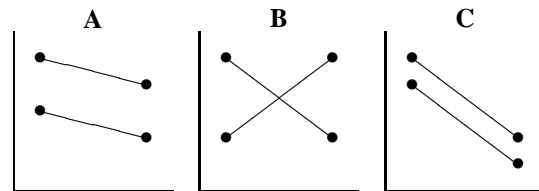
11. Refer to the table given in the previous question.

Use the standard normal curve to obtain an approximate P-value for the MH test and the first alternative.

12. A researcher decides to perform Fisher's test with the first alternative on the collapsed table. Use the standard normal curve to obtain an approximate P-value. (Hint: Forty-four of the 100

females and 28 of the 100 males in the sample yielded successes.)

13. Three interaction graphs, A, B, and C, are partially drawn below. (Partial because the values for the collapsed table are missing.)



Which one or more of these interaction graphs could possibly provide an example of Simpson's paradox? (In other words, for which of these interaction graphs is it mathematically possible to have values of \hat{p} 's and \hat{q} 's for the collapsed table so that Simpson's paradox occurs?)

Explain your answer.

14. A researcher has independent random samples of 100 females and 100 males. Each subject yields a response of success or failure. The collapsed table of 200 individuals is divided into three component tables on the basis of a factor that has possible values low, moderate, and high. The component tables are given below.

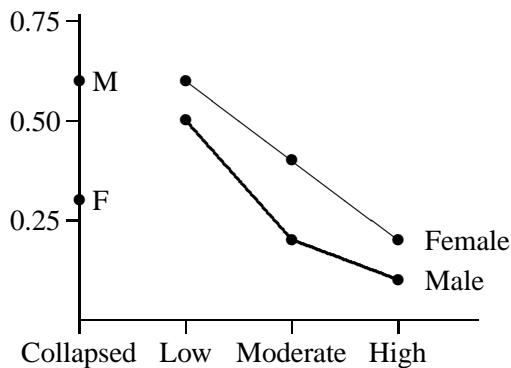
	Low		Total
	S	F	
Female	30	20	50
Male	24	16	40
Total	54	36	90

	Moderate		Total
	S	F	
Female	12	18	30
Male	4	6	10
Total	16	24	40

	High		Total
	S	F	
Female	6	14	20
Male	15	35	50
Total	21	49	70

Compute the female weights, as described in the section on standardized rates in the text. Be careful to label your weights (that is, tell me what each weight represents).

15. Explain why the following picture *cannot* be an interaction graph for any data set.



16. Compute the weighted average of the two numbers, 0.4 and 0.8, using the weights 0.25 and 0.75, respectively.
17. Which of the following sets of numbers do not satisfy the conditions for being weights? (Choose one or more.)
- $w_1 = 0, w_2 = 0, w_3 = 1$
 - $w_1 = 0.2, w_2 = 0.3, w_3 = 0.4$
 - $w_1 = 0.6, w_2 = 0.5, w_3 = -0.1$
 - $w_1 = 0.6, w_2 = 0.3, w_3 = 0.1$

9.4 Solutions to Exam Questions

- 60.
- The answer to question 2 is given in the statement of question 7.
- (a).
- (b).
- (d).
- (c).
- False.
- True.
- True.
- $z = 0.77$, and the approximate P-value is 0.2206.

11. $z = 1.37$, and the approximate P-value is 0.0853.

12. $z = 2.35$, and the approximate P-value is 0.0094.

13. C.

Simpson's paradox cannot occur for B because there is not a consistent pattern in the component tables.

Picture A is a bit trickier. Suppose, without lack of generality, that the vertical axis plots \hat{p} and that the higher line segment is from population one and the lower line segment is from population two. There is a consistent direction in the component tables, with $\hat{p}_1 > \hat{p}_2$ in each table. Notice, however, that the smaller value of \hat{p}_1 is larger than the larger value of \hat{p}_2 ; in words, there is no "overlap." It now follows that \hat{p}_1 for the collapsed table *must* be larger than \hat{p}_2 for the collapsed table. Thus, the reversal needed for Simpson's paradox is not possible.

For picture C there is overlap of the values of \hat{p} , making the reversal necessary for Simpson's paradox *possible*.

14. The female weights are $w_{1L} = 0.50$, $w_{1M} = 0.30$, and $w_{1H} = 0.20$.

15. The proportion of successes for males in the the collapsed table is a weighted average of the proportions of successes for males in the component tables. Thus, it cannot be larger than all of the latter numbers as is pictured.

16. 0.7.

17. b and c cannot be weights.