

STAT 854. Biometrics: Methods in Biomedical Studies (Spring 2023)

Homework 2: due on Friday Mar 17 2023 by 5:00pm via the course canvas page

Problem 1. Give two practical examples of incomplete data:

- (i) Describe the examples using the language that non-statisticians can understand.
- (ii) Formulate the settings to present their differences and similarities to the well-known incomplete data structures.

You may choose to use the information from an accessible journal or website. If so, provide the full citation.

Problem 2. The data summarized in Table 1 come from a prospective cohort study. You may suppose the cohort is closed and that there are no missing data.

- Use the data to study the association between serum cholesterol level and coronary heart disease (CHD). Subject age and sex are provided as they may confound/modify this association.

Write up the results of your investigation as for the Results section of a published paper. FYI, The STROBE guidelines (Vandenbroucke et al. 2007), von Elm et al. (2007)) might be helpful in deciding what material is appropriate for the Results section of a paper.

Table 1: Proportions of CHD cases

Sex	Age Group (years)	Serum Cholesterol Level (mg/100 ml)			
		<190	190-219	220-249	≥ 250
Male	30-49	13/340	18/408	40/421	57/362
	50-62	13/123	33/176	35/174	49/183
Female	30-49	6/542	5/552	10/412	18/357
	50-62	9/58	12/135	21/218	48/395

Problem 3. Rothman and Greenland (2005) is an influential paper on causal inference.

- (i) Provide a brief description of Rothman and Greenland’s sufficient component cause model.
- (ii) Use Figure 1 from Rothman and Greenland (2005) to answer the following questions. Provide a brief explanation for your answer.
 - Are any of the component causes displayed in Figure 1 necessary?
 - Suppose that the three diagrams in Figure 1 each account for one third of disease incidence. (i.e. there are only three combinations of component causes resulting in disease) What would be the consequence on disease incidence of eliminating component cause B?
 - Using Figure 1 as a guide draw all possible schematic diagrams of sufficient causes for Tay-Sachs disease.
- (iii) Doll and Peto (1981) estimated the proportion of deaths in the United States attributed to various factors (see Table 2). Provide a critique of this table based on the sufficient component cause model described by Rothman and Greenland (2005).

Table 2: Percentage of cancer deaths in the USA attributed to different factors

Factor	% of U.S.A cancer deaths
Diet	35
Tobacco	30
Infection	10
Reprod. sexual behavior	7
Occupation	4
Alcohol	3
Geophysical (e.g. U.V.)	3
Pollution	2
Medicines/medical procedures	1
Food additives	< 1
Industrial products	< 1
Total	≈ 97

Problem 4. Breast cancer mortality rates are provided in Table 3 by marital status (Colton, 1974 pp. 295-296).

Table 3: Breast cancer mortality rates by marital status, US females, 1929-1931.

Marital Status	Deaths per 100,000 Women	
	Crude	Adjusted*
Single	15.0	43.7
Married	24.5	25.0
Divorced	29.3	35.7
Widowed	74.4	28.3

* Age adjusted to total US female population in 1930

(i) State your findings from the table and explain them briefly.

(ii) Suppose you are asked to provide a prediction for adjusted death rates of the four groups in 2017 using the information given in Table 3. Describe what additional information or assumptions you need and your procedure to obtain the prediction.

References

1. Colton T. Statistics in Medicine. Boston, MA: Little, Brown and Company 1974.
2. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. Journal of the National Cancer Institute 1981;66:1191-1308.
3. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. American Journal of Public Health 2005 (Supplement 1):95(s1):s144-s150.
4. Vandembroucke JP, von Elm E, Altman DG et al. STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. Annals of Internal Medicine 2007; 147(8):W163- W194.
5. von Elm E, Altman DG, Egger M, et al. STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. British Medical Journal 2007; 335(7624):806-808. (<http://www.strobe-statement.org/>)