MATH 3316: Assignment 3

Due Time/Date: 23:59 Sep. 18, 2025

Please submit your assignment online in .pdf format. Please do not send the assignment by email or hand in the assignment in person.

1. (IPS10-2.146) More smokers live at least 20 more years! You can see the headlines: “More smokers than nonsmokers live at least 20 more years after being contacted for study!” A medical study contacted randomly chosen people in a district in England. Here are data on the 1314 women contacted who were either current smokers or who had never smoked. The tables classify these women by their smoking status and age at the time of the survey and whether they were still alive 20 years later. (15 Points)

Age 18 to 44 Age 45 to 64 Age 65+

Smoker Not Smoker Not Smoker Not

|  |  |  |  |
| --- | --- | --- | --- |
| Dead | 19 13 | 78 52 | 42 165 |
| Alive | 269 327 | 167 147 | 7 28 |

* 1. From these data, make a two-way table of smoking (yes or no) by dead or alive. What percent of the smokers stayed alive for 20 years? What percent of the nonsmokers survived? It seems surprising that a higher percent of smokers stayed alive.
  2. The age of the women at the time of the study is a lurking variable. Show that within each of the three age groups in the data, a higher percent of nonsmokers remained alive 20 years later. This is another example of Simpson’s paradox.
  3. The study authors give this explanation: “Few of the older women (over 65 at the original survey) were smokers, but many of them had died by the time of follow-up.” Compare the percent of smokers in the three age groups to verify the explanation.

1. (IPS10-2.100) Survival and class on the Titanic. On April 15, 1912, on her maiden voyage, the Titanic collided with an iceberg and sank. The ship was luxurious but did not have enough lifeboats for the 2224 passengers and crew. As a result of the collision, 1502 people died.28 The level of luxury and the price of the ticket varied with the class, first class being the most luxurious. There were 323 passengers in first class, 277 in second class, and 709 in third class. The number of first-class passengers who survived was 200. For second- and third-class passengers who survived, the numbers were 119 and 181, respectively. Let’s look at these data with a two-way table.(25 Points)

[ex02-100titanic.csv](https://media.saplinglearning.com/priv/he/stats/ips10e/datafiles/chapter_02/csv/ex02-100titanic.csv)

* 1. Create a two-way table that you could use to explore the relationship between survival and class.
  2. Which variable is the explanatory variable, and which is the response variable? Give reasons for your answers.
  3. Find the two marginal distributions. Write a brief summary explaining the major features of these distributions.
  4. Which conditional distribution would you choose to explain the relationship between these two variables? Explain your answer.
  5. Which conditional distribution would you choose to explain the relationship between these two variables? Explain your answer.

1. (IPS10-2.105) Which hospital is safer? Insurance companies and consumers are interested in the performance of hospitals. The government releases data about patient outcomes in hospitals that can be useful in making informed health care decisions. Here is a two-way table of data on the survival of patients after surgery in two hospitals. All patients undergoing surgery in a recent time period are included. “Survived” means that the patient lived at least six weeks following surgery. (10 Points)

Hospital A Hospital B Died 63 16

Survived 2037 784

Total 2100 800

What percent of Hospital A patients died? What percent of Hospital B patients died? These are the numbers one might see reported in the media.

1. (IPS10-2.106) Patients in “poor” or “good” condition. Refer to the previous exercise. Not all surgery cases are equally serious. Patients are classified as b eing in e ither “ poor” or “good” condition before surgery. Here are the data broken down by patient condition. The entries in the original two-way table are just the sums of the “poor” and “good” entries in this pair of tables. (20 Points)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Good Condition** | Hospital A | Hospital B |  |
|  | Died | 6 | 8 |  |
|  | Survived | 594 | 592 |  |
|  | Total | 600 | 600 |  |
|  |  |  |  |  |
|  | **Bad Condition** | Hospital A | Hospital B |  |
|  | Died | 57 | 8 |  |
|  | Survived | 1443 | 192 |  |
|  | Total | 1500 | 200 |  |

* 1. Find the death rate for Hospital A patients who were classified as “poor” before surgery. Do the same for Hospital B. In which hospital do “poor” patients fare better?
  2. Repeat part (a) for patients classified as “good” before surgery.
  3. What is your recommendation to someone facing surgery and choosing between these two hospitals?
  4. How can Hospital A do better in both groups, yet do worse overall? Look at the data and carefully explain how this can happen.

1. (IPS10-2.114) Marriage and income. Data show that men who are married, and also divorced or widowed men, earn quite a bit more than men who have never been married. This does not mean that a man can raise his income by getting married. Suggest several lurking variables that you think are confounded with marital status and that help explain the association between marital status and income. Use a diagram like Figure 2.34(c) [the figure in the slides] to illustrate your ideas. (15 Points)
2. (IPS10-2.116) Computer chip manufacturing and miscarriages. A study showed that women who work in the production of computer chips have abnormally high numbers of miscarriages. The union claimed that exposure to chemicals used in production caused the miscarriages. Another possible explanation is that these workers spend most of their work time standing up. Illustrate these relationships in a diagram like one of those in Figure 2.34 [the figure in the slides]. (15 Points)