

Instructions

You can submit the homework either in paper or online

- Online: Take pictures of the written (theory) part; submit it (along with the simulation part) through Canvas
- Paper: Print out the result of the simulation part and staple it with the written work; hand it in at the beginning of the lecture on Friday (10/07)

1 Theory

- Problem 1: At a particular university, $1/2$ of the students drink alcohol and $1/3$ of the students smoke cigarettes. If we know further that $1/3$ of the students do neither, what fraction of the students does both?
- Problem 2: Event A has $P(A) = 0.7$. Event B has $P(B) = 0.3$. These events are independent. What is $P(A \cup B)$?
- Problem 3: In a community, 32% of the population smokers who like math; 27% are smokers who don't. What percentage of the population of this community smoke?
- Problem 4: From the three basic properties of probability (Useful Facts 3.1 in the text), prove that: If $A \subset B$, then $P(A) \leq P(B)$.

2 Simulations

- Problem 5:

In the library MASS, there is a dataset *UScereal* which contains information about popular breakfast cereals.

 - Produce a side-by-side bar plot to visualize the relationship between manufacturer (mfr) and shelf
 - Visualize the relationship between sodium and sugars
- Problem 6:

Load the built-in dataset *mtcars* (Motor Trend Car Road Tests). The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

- Produce a stacked bar plot to visualize the relationship between transmission (*am*) and cylinders (*cyl*)
- Make a scatterplot of cylinders (*cyl*) vs. miles per gallon (*mpg*).

For each problem, attach the script (.R) containing the commands, and a screenshot of the result of the script. The plots should have clear titles and all axes labeled.