Data 100 **(Principles and Techniques of Data Science)**

Disc 2 Bonus Questions Solutions

Term: **Spring 2020** *Name:*

Problem Bonus - from Short Quiz

Consider a group of 20 students at Berkeley:

- 2 of them are currently freshmen.
- 12 of them are currently sophomores.
- 6 of them are currently juniors.

Part I

If we draw a simple random sample of 5 students and **only require 3 of them to be sophomores**, while the rest of the 2 can be both freshmen, both juniors, or one of each, what is the probability of drawing such a sample?
Solution:
We can rephrase the question as choosing 3 sophomores and 2 **non-sophomores**.
As we can see, there are 12 sophomores in total, and we want to choose 3 of them. This gives us a total of $\binom{12}{3}$ ways.
There are $20 - 12 = 8$ non-sophomores in total, and we want to choose 2 of them. This gives us a total of $\binom{8}{2}$.
Hence, the total number of ways of choosing 3 sophomores and 2 non-sophomores is equal to:

$$\begin{aligned} &\#\text{ ways choosing 3S and 2 non-S} = \#\text{ ways of choosing 3S} \times \#\text{ ways of choosing 2 non-S} \\ &= \binom{12}{3} \times \binom{8}{2} \end{aligned}$$

Recall from discussion, since we also have to consider how many ways there are to choose **any 5 students** from the overall population of 20, this gives us $\binom{20}{5}$ ways.

Hence,

$$P(\text{3S and 2 non-S}) = \frac{\binom{12}{3}\binom{8}{2}}{\binom{20}{5}}.$$

Part II

Same setting, but this time, if we want to draw a simple random sample that contains **3 students from 1 particular grade** and **2 students from a different grade**, what is the probability of drawing such a sample?
Solution:
This question is a bit tricky now since we never specify the groups to draw the 2/3 students from and not every group can support a draw of 3 students (i.e. freshmen).
A good way to solve this proble is by explicitly enumerating the possible ways of choosing the grades of students.

Suppose the ordered pair (A, B) represents the two grades we choose, and:

- We will draw 3 students from group A.
- We will draw 2 students from group B.

We can **list out all the ways of choosing groups A and B**.

- 3 Sophomores, 2 Freshmen
- 3 Sophomores, 2 Juniors
- 3 Juniors, 2 Freshmen
- 3 Juniors, 2 Sophomores

We can then calculate individually (for each of those cases), the number of ways of choosing 2/3 students respectively from the grades. Specifically speaking,

$$\begin{aligned}
\#\text{ ways of 3 Sophomores, 2 Freshmen} &= \binom{12}{3} \times \binom{2}{2} \\
\#\text{ ways of 3 Sophomores, 2 Juniors} &= \binom{12}{3} \times \binom{6}{2} \\
\#\text{ ways of 3 Juniors, 2 Freshmen} &= \binom{6}{3} \times \binom{2}{2} \\
\#\text{ ways of 3 Juniors, 2 Sophomores} &= \binom{6}{3} \times \binom{12}{2}
\end{aligned}$$

Again, the denominator of this probability we are computing will be the chance of choosing any 5 students from a population of 20, which will be $\binom{20}{5}$ ways.

Hence, summing up the total number of ways of assigning grades of students, we have our probability to be:

$$P(\text{3A, 2B}) = \frac{\binom{12}{3}\binom{2}{2} + \binom{12}{3}\binom{6}{2} + \binom{6}{3}\binom{2}{2} + \binom{6}{3}\binom{12}{2}}{\binom{20}{5}}$$