AML5103 | Applied Probability and Statistics | Sessional-1 Solutions

1. [10 points] [CO 1, BT 3] There are 30 students (15 males and 15 females) participating in a study at MSIS. For the study, 10 students are randomly selected by a computer and are assigned places in a line where ordering matters. Based on this information, please answer the following questions (you need not simplify the final answer):

   (a) The computer is malfunctioning, and it will always select Sudhanva, Keerthana, and Nidhi to be included in the group of 10 selected students. It will also randomly select 7 students to make up a group of 10. The computer also always places Sudhanva, Keerthana, and Nidhi together in the line with any ordering of the three being possible. How many possible line configurations can the computer produce? You may answer the question by filling in the blanks below:

   There are ____ spots in the line to put the group of ____ and ____ ways to arrange them. The remaining ____ spots are filled with ____ other students so we have ____ ways to choose the 7 students and ____ ways to arrange them in the remaining ____ spots.

   Solution: There are <u>8</u> spots in the line to put the group of <u>3</u> and <u>3!</u> ways to arrange them. The remaining <u>7</u> spots are filled with <u>27</u> other students so we have $\underline{27C_7}$ ways to choose the 7 students and 7! ways to arrange them in the remaining 7 spots. Therefore, the final answer is $8C_1 \times 3! \times \times 27C_7 \times 7!$.

(b) The computer is still malfunctioning: again it still randomly selects 7 students to go with Sudhanva, Keerthana, and Nidhi. However, these three students no longer have to be grouped together in the line, but Sudhanva must be in front of Keerthana and Keerthana must be in front of Nidhi. How many possible line configurations can the computer produce?

Solution: We can choose the 3 spots for the students in the line in $10C_3$ ways. The remaining 7 spots are filled with 7 other students, so we have $27C_7$ ways to choose the 7 students and 7! ways to arrange them in the remaining 7 spots. Therefore, the final answer is $10C_3 \times 27C_7 \times 7!$.

2. [10 points] [CO 1, BT 3] One hundred students, including sixty men and forty women, are taking the APS course. The classroom has five rows of seats with twenty seats each. Assume all students attend every lecture but they are equally likely to sit anywhere each time.

(a) What is the probability there are exactly eight women in the first row for a given lecture?

Solution: The sample space

$$S = \{(m_1, m_2, m_3, m_4, \ldots, m_{20}), (m_1, f_1, m_2, f_2, \ldots, m_{10}, f_{10}), \ldots, (f_{31}, f_{32}, f_{33}, \ldots, f_{40})\}.$$

The sampling space is
$$s = \{m_1, \dots, m_{60}, f_1, \dots, f_{40}\}.$$
The out comes in the sample space are equally likely with each outcome having a probability measure $1/n(S)$ where $n(s)$ is the number of outcomes in the sample space which is equal to the number of ways to select 20 out of 100 objects from the sampling space s without replacement and order does not matter $= 100C_{20}$. The event E is the set

$$\left\{\left(\underbrace{f_1, f_2, \dots, f_8}_{\text{8 women}}, \underbrace{m_1, m_2, \dots, m_{12}}_{\text{12 men}}\right), \dots, \left(\underbrace{m_{49}, m_{50}, \dots, m_{60}}_{\text{12 men}}, \dots, \underbrace{f_{33}, f_{34}, \dots, f_{40}}_{\text{8 women}}\right)\right\}.$$

The probability of event E is $n(E)/100C_{20} = (40C_8 \times 60C_{12})/100C_{20}$ where the numerator corresponds to selecting 8 out 40 women and 12 out of 60 men both without replacement and order does not matter.

(b) What is the probability there is at least one woman in the first row for a given lecture?

Solution: We use $P(E) = 1 - P(E^c)$, where E^c is the event corresponding to no woman which is the set
$$\{(m_1, m_2, \dots, m_{20}), \dots, (m_{49}, m_{50}, \dots, m_{60})\}.$$
The probability of E^c is $60C_{20}/100C_{20}$ where the numerator corresponds to selecting 20 out 60 men without replacement and order does not matter $\Rightarrow P(E) = 1 - 60C_{20}/100C_{20}$.

3. [10 points] [CO 1, BT 3] We have a box with 20 balls in four different colors. Out of the 20 balls, there are 5 balls for each color. Balls for each color are numbered from 1 to 5. We pick 4 balls from the box *without replacement*. Please answer the following questions:

   (a) What is the probability that we get two different colors, two balls each?

   Hint: the number of favorable outcomes can be calculated as the number of ways to select ____ colors out of ____ and select ____ balls from ____ of the first color chosen and select ____ balls from ____ of the second color chosen.

Solution: The sample space is

$$S = \{(r_1, r_2, r_3, r_4), (r_1, g_1, b_1, y_1), (r_2, r_3, r_4, r_5), \ldots, (b_2, b_3, b_4, b_5)\}$$

and the sampling space is

$$s = \{r_1, ..., r_5, g_1, ..., g_5, b_1, ..., b_5, y_1, ..., y_5\}\,.$$

Outcomes in the sample space are equally likely with probability measure equal to $1/n(S)$, where $n(S)$ is the number of outcomes in the sample space which is equal to the number of ways to select 4 out 20 objects without replacement and order does not matter $= 20C_4$. Event E is the set

$$\{(r_1, b_1, r_2, b_2), (g_1, g_2, y_1, y_2), ..., (b_1, b_2, y_1, y_2)\}\,.$$

The probability of event E is $(1/20C_4) \times n(E) = (4C_2 \times 5C_2 \times 5C_2)/(20C_4)$, where the numerator corresponds to the number of ways to select $\underline{2}$ colors out of $\underline{4}$ and select $\underline{2}$ balls from $\underline{5}$ of the first color chosen and select $\underline{2}$ balls from $\underline{5}$ of the second color chosen.

(b) What is the probability that we get four different colors and sum of the numbers on the balls is equal to 8?

Hint: for the number of favorable outcomes, think of 8 identical objects and 4 distinct non-empty bins.

Solution: The event E is the set

$$\{(r_1, b_2, g_3, y_2), (r_2, b_2, g_2, y_2), \ldots, (r_5, b_1, g_1, y_1)\}.$$

The probability of event E is $(1/20C_4) \times n(E)$, where we calculate $n(E)$, the number of outcomes in E, as follows: we first start with the arrangement $|1|1|1|1|$ which has the number 1 in each of the 4 distinct bins. This means, the ball number for a particular arrangement of the colors should be at least 1; for example, (r_1, g_1, b_1, y_1). Then, we look at distributing 4 identical objects to the 4 distinct bins which already have 1 object each. For example, one way of doing it is $|4|0|0|0$ which when added to the previous one-object-in-each-bin scenario results in $|5|1|1|1|$. This corresponds to the outcome (r_5, b_1, g_1, y_1). So, $n(E)$ is equal to the number of ways to distribute 4 identical objects to 4 distinct bins which is equal to $(4 + 4 - 1)C_4 = 4! \times 7C_4$. Therefore, $P(E) = (7C_4)/(20C_4)$.

4. [10 points] [CO 2, BT 4] Suppose that an insurance company classifies people who buy two-wheeler insurance from them into one of four classes: good, fair, average, and bad risks. As a data scientist for the company, you have access to the following customer data for the calendar year 2021-22:

Class	Total no. of riders	No. involved in accident
Bad risk	5000	2000
Average risk	4000	1200
Fair risk	3000	600
Good risk	3000	300

(a) What is the probability that a new customer will meet with an accident during 2022-23?

Solution: Using the law of total probability, we get

$$P\left(\text{Accident}\right) = P\begin{pmatrix} \text{Accident AND Bad} \\ OR \\ \text{Accident AND Avg.} \\ OR \\ \text{Accident AND Fair} \\ OR \\ \text{Accident AND Good} \end{pmatrix} = \begin{matrix} P\left(\text{Accident AND Bad}\right) \\ + \\ P\left(\text{Accident AND Avg.}\right) \\ + \\ P\left(\text{Accident AND Fair}\right) \\ + \\ P\left(\text{Accident AND Good}\right) \end{matrix}$$

$$= \begin{matrix} P\left(\text{Accident} \,|\, \text{Bad}\right) \times P\left(\text{Bad}\right) \\ + \\ P\left(\text{Accident} \,|\, \text{Avg.}\right) \times P\left(\text{Avg.}\right) \\ + \\ P\left(\text{Accident} \,|\, \text{Fair}\right) \times P\left(\text{Fair}\right) \\ + \\ P\left(\text{Accident} \,|\, \text{Good}\right) \times P\left(\text{Good}\right) \end{matrix} = \begin{matrix} \frac{2000}{5000} \times \frac{5000}{15000} \\ + \\ \frac{1200}{4000} \times \frac{4000}{15000} \\ + \\ \frac{600}{3000} \times \frac{3000}{15000} \\ + \\ \frac{300}{3000} \times \frac{3000}{15000} \end{matrix} \approx 0.27.$$

(b) If a customer A had not met with an accident during 2021-22, what is the probability that he or she is not a good rider?

Solution:

$$
\begin{aligned}
P\left(\text{Good rider}^c \,|\, \text{Accident}^c\right) &= 1 - P\left(\text{Good rider} \,|\, \text{Accident}^c\right) \\
&= 1 - \frac{\left[1 - \underbrace{P\left(\text{Accident} \,|\, \text{Good rider}\right)}_{=300/3000}\right] \times \underbrace{P(\text{Good rider})}_{=3000/15000}}{\underbrace{P\left(\text{Accident}^c\right)}_{=1-P(\text{Accident})=1\text{-}0.27}} \\
&\approx 0.75.
\end{aligned}
$$

5. [10 points] [CO 2, BT 4] You have tracked the performance of the local meteorologist and compiled the following data:

$$
\begin{aligned}
\text{P(forecast rain } \underline{\text{and}} \text{ actual rain)} &= 0.45, \\
\text{P(forecast no rain } \underline{\text{and}} \text{ actual rain)} &= 0.05, \\
\text{P(forecast no rain } \underline{\text{and}} \text{ actual no rain)} &= 0.35.
\end{aligned}
$$

(a) How often does she forecast rain?

Solution:

$$P(\text{Forecast rain}) = \underbrace{P(\text{Forecast rain}\,\underline{\text{AND}}\,\text{Rain})}_{=0.45} + \underbrace{P(\text{Forecast rain}\,\underline{\text{AND}}\,\text{No rain})}_{=1-(0.45+0.05+0.35)} = 0.6.$$

(b) How often does she not make a mistake?

Solution:

$$\begin{aligned} P(\text{No mistake}) &= P(\text{Forecast rain}\,\underline{\text{AND}}\,\text{Rain}\ \boxed{\text{OR}}\ \text{Forecast no rain}\,\underline{\text{AND}}\,\text{No rain}) \\ &= \underbrace{P(\text{Forecast rain}\,\underline{\text{AND}}\,\text{Rain})}_{=0.45} + \underbrace{P(\text{Forecast no rain}\,\underline{\text{AND}}\,\text{No rain})}_{=0.35} = 0.8. \end{aligned}$$

(c) Given that it did not rain today, what is the probability that she forecast rain in last night's broadcast?

Solution:

$$P\left(\text{Forecast rain} \,|\, \text{No rain}\right) = \frac{P(\text{Forecast rain}\,\underline{\text{AND}}\,\text{No rain})}{P(\text{No rain})} = \frac{1-(0.45+0.05+0.35)}{0.35+0.15} = 0.3.$$