Probability and Statistics for JMC
Solutions 2a — Probability

1. A string of 4 letters is generated at random. Let E_j denote the event that the j^{th} letter is L.

   (a) Describe the following events in words?

      i. $E_1 \cap E_2 \cap E_3 \cap E_4 =$ the string is LLLL
      ii. $E_1 \cup E_2 \cup E_3 \cup E_4 =$ the string has at least one L
      iii. $\overline{E_1} \cap \overline{E_2} \cap \overline{E_3} \cap \overline{E_4} =$ the string has no L's
      iv. $E_3 \cap E_4 =$ the last two letters are LL (i.e. the string is ??LL, where ? can be any letter)
      v. $E_1 \cap \overline{E_2} =$ the first letter is L and the second letter is definitely not an L (i.e. L*??, where * is not an L and ? can be anything)

   (b) Write these events in set notation

      i. L appears at least 3 times $= (E_1 \cap E_2 \cap E_3) \cup (E_1 \cap E_2 \cap E_4) \cup (E_1 \cap E_3 \cap E_4) \cup (E_2 \cap E_3 \cap E_4)$
      ii. L appears exactly once $= (E_1 \cap \overline{E_2} \cap \overline{E_3} \cap \overline{E_4}) \cup (\overline{E_1} \cap E_2 \cap \overline{E_3} \cap \overline{E_4}) \cup (\overline{E_1} \cap \overline{E_2} \cap E_3 \cap \overline{E_4}) \cup (\overline{E_1} \cap \overline{E_2} \cap \overline{E_3} \cap E_4)$
      iii. We do not have two L's in a row $= \overline{(E_1 \cap E_2) \cup (E_2 \cap E_3) \cup (E_3 \cap E_4)}$

2. Let S be a set and $\mathcal{L}$ be a set of σ-algebras on S. Define

$$\mathcal{E}^* = \{A \subseteq S : A \in \mathcal{F} \text{ for all } \mathcal{F} \in \mathcal{L}\}.$$

   Show that $\mathcal{E}^*$ is a σ-algebra.

   We need to show $\mathcal{E}^*$ obeys the three properties that define a σ-algebra.

   (1) S is a member of every $\mathcal{F}$ in $\mathcal{L}$ since each $\mathcal{F}$ is a σ-algebra. Therefore $S \in \mathcal{E}^*$.
   (2) Consider any $A \in \mathcal{E}^*$. By definition of $\mathcal{E}^*$, $A \in \mathcal{F}$ for every $\mathcal{F}$ in $\mathcal{L}$. Then, since each $\mathcal{F}$ is itself a σ-algebra on S, $\overline{A}$ must be a member of each $\mathcal{F}$. Then, by definition of $\mathcal{E}^*$, $\overline{A} \in \mathcal{E}^*$.
   (3) Consider $A_1, A_2, \ldots$ which are all elements of $\mathcal{E}^*$. Then A_1 must be an element of all the $\mathcal{F}$'s in $\mathcal{L}$, and similarly for $A_2, A_3, \ldots$. But since each $\mathcal{F}$ is a σ-algebra that contains all the A_i's, each $\mathcal{F}$ must contain the union of the A_i's. The union of the A_i's is in every $\mathcal{F}$ and therefore also in $\mathcal{E}^*$.

3. Let P be a probability measure, S be a sample space and A_1, A_2, A_3 be events (subsets of S). Show that

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2 \cup A_3) - P((A_1 \cap A_2) \cup (A_1 \cap A_3)) \\ &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_2 \cap A_3) \\ &\quad - P(A_1 \cap A_3) + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

Use the identity that $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F) - \mathrm{P}(E \cap F)$ for any events E and F. Write $A_1 \cup A_3 \cup A_3$ as the union of two sets A_1 and $(A_2 \cup A_3)$ and then apply the above identity. Use the distributive property of set intersection in the $\mathrm{P}(A_1 \cap (A_2 \cup A_3))$ term to get the first equality. Then use the identity again to expand out the $\mathrm{P}(A_2 \cup A_3)$ and $\mathrm{P}((A_1 \cap A_2) \cup (A_1 \cap A_3))$ terms to get the second equality.

4. Suppose two events E and F are mutually exclusive. State the precise conditions under which they may also be independent.

If E and F are mutually exclusive it means they cannot both simultaneously occur, i.e. there is no outcome in their intersection: $E \cap F = \emptyset$. That means that $\mathrm{P}(E \cap F) = \mathrm{P}(\emptyset) = 0$. Independence means that $\mathrm{P}(E \cap F) = \mathrm{P}(E)\mathrm{P}(F)$. Therefore, it they are mutually exclusive and independent we have $0 = \mathrm{P}(E)\mathrm{P}(F)$, which implies that at least one of E or F has probability 0. In words, if two events are mutually exclusive and have non-zero probability of occurring, then they cannot be independent.

5. What is the probability that a single roll of a die will give an odd number if

   (a) no other information is given;

   $\mathrm{P}(\text{odd}) = \mathrm{P}(\{1, 3, 5\}) = \frac{|\{1,3,5\}|}{|\{1,2,3,4,5,6\}|} = \frac{1}{2}$.

   (b) you are told that the number is less than 4.

   Two ways. With conditional probability arithmetic:
   $\mathrm{P}(\text{odd} \mid \text{less than } 4) = \mathrm{P}(\text{odd and less than } 4)/\mathrm{P}(\text{less than } 4) = \frac{2/6}{3/6} = \frac{2}{3}$.
   Restricting sample space to outcomes less than 4:
   $\mathrm{P}(\text{odd} \mid \text{less than } 4) = \frac{|\{1,3\}|}{|\{1,2,3\}|} = \frac{2}{3}$.

6. (a) What's the probability of getting two sixes with two dice?

   There are 36 possibilities only one of which is two sixes so answer is $\frac{1}{36}$.
   Or use independence of two rolls:
   $\mathrm{P}(6 \text{ on first } \cap \ 6 \text{ on second}) = \mathrm{P}(6 \text{ on first})\mathrm{P}(6 \text{ on second}) = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}$.

   (b) What's the probability of getting a total of 3 with two dice?

   There are 2 out of the 36 outcomes where the total is 3: $\{(1, 2), (2, 1)\}$. So probability is $\frac{2}{36} = \frac{1}{18}$.

7. Two students try to solve a problem they've been set. Student A has a probability of $\frac{2}{5}$ of being able to solve the problem, and student B has a probability of $\frac{1}{3}$. If both try it independently, what is the probability that the problem is solved?

Two ways.
P(solved) $=$ 1 $-$ P(not solved) $=$ 1 $-$ P(not solved by A and not solved by B) $=$ $1-$ P(not solved by A)P(not solved by B) $= 1-(1-\frac{2}{5})(1-\frac{1}{3}) = 1-(\frac{3}{5})(\frac{2}{3}) = \frac{3}{5}$.
Or use $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F) - \mathrm{P}(E \cap F)$.
P(solved) $=$ P(solved by A $\cup$ solved by B) $=$ P(solved by A) $+$ P(solved by B) $-$ P(solved by A and solved by B) $= \frac{2}{5}+\frac{1}{3}-(\frac{2}{5})(\frac{1}{3}) = \frac{6}{15}+\frac{5}{15}-\frac{2}{15} = \frac{3}{5}$.

8. A straight line AB of unit length is divided internally at a point X, where X is equally likely to be found anywhere along AB. What is the probability that $|AX||XB| < \frac{3}{16}$? ($|AX|$ is the length of the line segment from A to the internal point X and similarly for $|XB|$.)

Call the length of the left piece x (and x is constrained to be between 0 and 1). Then $|AX||XB| = x(1-x)$ and the inequality becomes $x(1-x) < \frac{3}{16}$. so the inequality is satisfied when $x < \frac{1}{4}$ or $x > \frac{3}{4}$. Using the classical interpretation of probability, the probability that the inequality is satisfied is the "size of the region where x satisfies the inequality" divided by "size of the sample space". That is $\frac{(\frac{1}{4}-0)+(1-\frac{3}{4})}{1} = \frac{1}{2}$. Or, since the two regions are mutually exclusive,
$\mathrm{P}\left(|AX||XB| < \frac{3}{16}\right) = \mathrm{P}(x<\frac{1}{4}) + \mathrm{P}(x>\frac{3}{4}) = \frac{1}{4}+\frac{1}{4} = \frac{1}{2}$.

9. A fair die is thrown repeatedly until a six is obtained. Let A_k be the event that the first six occurs on the k^{th} throw. Assume that

$$P(A_k) = \frac{1}{6}\left(\frac{5}{6}\right)^{k-1} \qquad k = 1, 2, 3, \ldots$$

Let B be the event that an even number of throws are required to obtain a six. Find $P(B)$.

The event B is the union of the even A_i's: $B = A_2 \cup A_4 \cup A_6 \cup \ldots$. Note that the A_i's are a collection of pairwise disjoint events: there is no way that the first six appears on the i^{th} throw and that the first six also appears on the j^{th} throw (for $i \neq j$). Therefore,

$$\begin{aligned}\mathrm{P}(B) &= \mathrm{P}(A_2 \cup A_4 \cup A_6 \cup \ldots) \\ &= \mathrm{P}(A_2) + \mathrm{P}(A_4) + \mathrm{P}(A_6) + \ldots \\ &= \frac{1}{6}\left(\frac{5}{6}\right) + \frac{1}{6}\left(\frac{5}{6}\right)^3 + \frac{1}{6}\left(\frac{5}{6}\right)^5 + \ldots.\end{aligned}$$

This is the sum of geometric series with first term $\frac{1}{6}\left(\frac{5}{6}\right)$ and ratio between terms of $\left(\frac{5}{6}\right)^2$ (which is less than 1 so the sum converges). The result is

$$\mathrm{P}(B) = \frac{\frac{1}{6}\left(\frac{5}{6}\right)}{1-\left(\frac{5}{6}\right)^2} = \frac{5}{11}.$$

python simulation:

```
import numpy as np
#call toss(1) to return the number of tosses until the first six appears
toss = lambda n: n if np.random.rand()<1/6 else toss(n+1)
```

```
#simulate
k = np.array([toss(1) for i in range(100000)])
(k%2 == 0).sum()/len(k)
```

Output is `0.45576`.