Solutions to Problem Set 3

1. Let's start with the size of the sample space. Of the 5+4=9 balls, we choose 3 balls. So the denominator is $\binom{9}{3}$. Now,

$$P(\text{red is drawn more than once}) = 1 - P(\text{red is drawn zero times}) -$$

$$P(\text{red is drawn one time}).$$

If, for example, of the 3 balls we draw, 0 should be red, 3 balls are green. So, of the 5 red balls we select 0 and of the 4 green balls we select 3. The numerator of the corresponding probability is then $\binom{5}{0} \cdot \binom{4}{3}$.
Then the probability that a red ball is drawn more than once is given by

$$P(\text{red ball is drawn more than once}) = 1 - \frac{\binom{5}{0} \cdot \binom{4}{3}}{\binom{9}{3}} - \frac{\binom{5}{1} \cdot \binom{4}{2}}{\binom{9}{3}} = 0.595$$

2. Because two dices can have the same number of dots we are sampling with replacement.

   (a) After the first three dices are thrown the configuration for the second "draw" is fixed (because there exists only one combination that is equivalent to the first one). Therefore the probability is

   $$P(\text{same configuration}) = \frac{1}{n^k} = \frac{1}{6^3} = 0.0046$$

   (b) Although we can not distinguish the dice, we have to take permutations into account, because for example the combination $(1, 1, 1)$ is only possible once, $(1, 2, 3)$ in 3! ways, which means those events will occur with a different probability. Therefore we consider the following three parts:

      i. Let A_1 be the event that all face values of the first three dice are different:
      Regarding the first throw, we are sampling without replacement since we require that all numbers are different. The possible number of combinations is therefore given by $\frac{n!}{(n-k)!}$.
      Once the outcome of the first throw is fixed, we require the

second throw to show exactly the same numbers, but since we can not distinguish the dice we have 3! possibilities. Therefore the total number of positive elementary events in A_1 is given by $\underbrace{\frac{n!}{(n-k)!}}_{\text{1. throw}} \cdot \underbrace{3!}_{\text{2. throw}}$.

ii. Let A_2 be the event that exactly two face values of the first three dice are the same:
Regarding the first throw exactly one number has to be different from the other two which have to be the same. This means we have 6 possibilities for the single number and 5 possibilities for the two numbers (not 6 since the two have to be different from the other number). For each of those outcomes we have $\binom{3}{2}$ possibilities since we can not distinguish the dice. Once the first throw is fixed we have again $\binom{3}{2}$ positive outcomes for the second throw, because - once again - we can not distinguish the dice. Therefore the total number of positive elementary events in A_2 is given by $\underbrace{6 \cdot 5 \cdot \binom{3}{2}}_{\text{1. throw}} \cdot \underbrace{\binom{3}{2}}_{\text{2. throw}}$

iii. Let A_3 be the event that all of the first three dice show the same number: In this case we have 6 possible outcomes for the first throw and one possible positive outcome for the second throw, i.e. showing the same numbers as in the first.

As the events A_1, A_2 and A_3 are disjoint and build a partition of all possible outcomes, we find since n^k is the number of possible elementary events in throwing 6 dice in total

$$\begin{aligned} P(\text{same configuration}) &= P(A_1) + P(A_2) + P(A_3) \\ &= \frac{\frac{6!}{(6-3)!}3!}{6^6} + \frac{6 \cdot 5 \cdot \binom{3}{2} \cdot \binom{3}{2}}{6^6} + \frac{6 \cdot 1}{6^6} = 0.0213 \end{aligned}$$

3. There are 30 students in a statistics class, 9 of whom are female. Also, 12 of the students are computer science majors, 4 of which are female.

   (a) Fill the blanks in the following table:

   (b) Based on this sample, construct a table for probabilities:

   (c) Using this data, give the following probabilities: $P(F)$, $P(CS)$,

	Female	Not Female	
CS	4	8	12
Not CS	5	13	18
	9	21	30

	Female	Not Female	
CS	2/15	4/15	2/5
Not CS	1/6	13/30	3/5
	3/10	7/10	1

$P(F \cap CS)$, $P(F|CS)$, $P(\overline{CS}|\overline{F})$.

$$\begin{aligned} P(F) &= \frac{3}{10} \\ P(CS) &= \frac{2}{5} \\ P(F \cap CS) &= \frac{2}{15} \text{ or } P(F \cap CS) = P(F) \cdot P(CS|F) = \frac{3}{10} \cdot \frac{4}{9} = \frac{2}{15} \\ P(F|CS) &= \frac{4}{12} \text{ or } P(F|CS) = \frac{P(F \cap CS)}{P(CS)} = \frac{2/15}{2/5} = \frac{1}{3} \\ P(\overline{CS}|\overline{F}) &= \frac{13}{21} \text{ or } P(\overline{CS}|\overline{F}) = \frac{P(\overline{CS} \cap \overline{F})}{P(\overline{F})} = \frac{13/30}{7/10} = \frac{13}{21} \end{aligned}$$

4. Consider the framework of an early test for HIV antibodies known as the ELISA test. The probability of a positive outcome given the tested person has HIV is 97.7% and of a negative outcome given the person is healthy is 92.6%. A study found that the probability of HIV occurring among North American population is about 0.26%. What is the probability that a person actually has HIV given they've tested positive?
   Define as HIV the event that a person has the virus and $T_{+/-}$ as a positive or negative test result. Then we have the following probabilities

   $$P(T_+|HIV) = 0.977 \qquad P(T_-|\overline{HIV}) = 0.926 \qquad P(HIV) = 0.0026$$

   We are looking for $P(HIV|T_+)$:

   $$P(HIV|T_+) = \frac{P(HIV \cap T_+)}{P(T_+)} = \frac{P(T_+|HIV)P(HIV)}{P(T_+|HIV)P(HIV) + P(T_+|\overline{HIV})P(\overline{HIV})}.$$

   Note that HIV and $\overline{HIV}$ form a partition, and

   $$P(\overline{HIV}) = 1 - P(HIV) = 1 - 0.0026 = 0.9974.$$

Also, $T_-|\overline{HIV}$ and $T_+|\overline{HIV}$ are complements, hence

$$P(T_+|\overline{HIV}) = 1 - P(T_-|\overline{HIV}) = 1 - 0.926 = 0.074.$$

Finally,

$$P(HIV|T_+) = \frac{0.977 \cdot 0.0026}{0.977 \cdot 0.0026 + 0.074 \cdot 0.9974} = 0.033.$$

5. The following table represents the data on survivors and deceased passengers of the Titanic distributed among classes.

	Class			
	1^{st}	2^{nd}	3^{rd}	Crew
Survived	203	118	178	212
Deceased	122	167	528	673

(a) What is the probability that a random person selected from this sample was travelling 1^{st} class?
First, let's fill in the missing data

	Class				
	1^{st}	2^{nd}	3^{rd}	Crew	Total
Survived	203	118	178	212	711
Deceased	122	167	528	673	1490
Total	325	285	706	885	2201

Probability of a random person travelling 1^{st} class is

$$P(Cl_1) = \frac{325}{2201} = 0.148.$$

(b) What is the probability that a particular survivor travelled with the crew? 1^{st} class?

$$P(Cr|S) = \frac{212}{711} = 0.298$$

$$P(Cl_1|S) = \frac{203}{711} = 0.286$$

(c) Are the events of being deceased and having travelled 3^{rd} class stochastically independent?
We will need to check whether $P(D) \cdot P(Cl_3) = P(D \cap Cl_3)$.

$$P(D) = \frac{1490}{2201} = 0.677, \qquad P(Cl_3) = \frac{706}{2201} = 0.321$$

$$P(D \cap Cl_3) = \frac{528}{2201} = 0.240, \qquad P(D) \cdot P(Cl_3) = 0.677 \cdot 0.321 = 0.217 \neq 0.240$$

Hence, the events are not stochastically independent.