SM4202 Exercise 1

1. (a) In combinatorics, the *inclusion-exclusion principle* states that for finite sets $A_1, \ldots, A_n$,

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{1\le i<j\le n} |A_i \cap A_j| + \sum_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n-1}|A_1 \cap \cdots \cap A_n|$$

Use the above formula to deduce the probability of $\mathrm{P}(A_1 \cup A_2 \cup A_3)$.

Solution: The inclusion-exclusion principle gives us $|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$. Thus,

$$\mathrm{P}\left(\bigcup_{i=1}^{3} A_i\right) = \sum_{i=1}^{3} \mathrm{P}(A_i) - \sum_{1\le i<j\le 3} \mathrm{P}(A_i \cap A_j) + \mathrm{P}(A_1 \cap A_2 \cap A_3).$$

(b) X is a number chosen at random from $\{1, 2, ..., 1000000\}$, so that each number is equally likely. Find the probability that X is divisible by one or more of the numbers 4,10 or 25.

Solution: Let E_k be the event that a number X chosen at random is divisible by k. We have that $\mathrm{P}(E_k) = 1/k$ (try to convince yourself this). Note that $E_4 \cap E_{10} = E_{20}$, and similarly for other intersections (think about lowest common multiples). We are then interested in $\mathrm{P}(E_4 \cup E_{10} \cup E_{25})$. Computing the the three parts of the sum above, we have

$$\begin{aligned}\Sigma_1 &= \mathrm{P}(E_4) + \mathrm{P}(E_{10}) + \mathrm{P}(E_{25}) \\ &= 1/4 + 1/10 + 1/25 = 39/100\end{aligned}$$

$$\begin{aligned}\Sigma_2 &= \mathrm{P}(E_4 \cap E_{10}) + \mathrm{P}(E_4 \cap E_{25}) + \mathrm{P}(E_{10} \cap E_{25}) \\ &= 1/20 + 1/100 + 1/50 = 2/25\end{aligned}$$

$$\begin{aligned}\Sigma_3 &= \mathrm{P}(E_4 \cap E_{10} \cap E_{25}) \\ &= 1/100\end{aligned}$$

Therefore, $\mathrm{P}(E_4 \cup E_{10} \cup E_{25}) = \Sigma_1 - \Sigma_2 + \Sigma_3 = 32/100$.

2. A fair coin will be tossed twice, the number N of heads will be noted, and then the coin will be tossed N more times. Let X be the total number of heads obtained.

(a) Decide on a probability space Ω, and make a table with the heading ω, $\mathrm{P}(\omega)$, and $X(\omega)$.

Solution:

	ω	$\mathrm{P}(\omega)$	$X(\omega)$
1	(T,T)	1/4	0
2	(H,T,H)	1/8	2
3	(H,T,T)	1/8	1
4	(T,H,H)	1/8	2
5	(T,H,T)	1/8	1
6	(H,H,H,H)	1/16	4
7	(H,H,H,T)	1/16	3
8	(H,H,T,H)	1/16	3
9	(H,H,T,T)	1/16	2

(b) Calculate the expectation $\mathrm{E}(X)$.

Solution: The expectation is $\mathrm{E}(X) = 1/4\times 0+1/8\times 2+\cdots+1/16\times 2 = 1.5$.

3. In a multiple choice examination Freda knows the correct answer with probability p; otherwise she guesses by randomly selecting one of the m possible answers. Given that Freda correctly answers a question, what is the probability that she guessed it?

Solution: Define the events

$$K = \{\text{Freda knows the answer}\} \text{ and } R = \{\text{Freda answers correctly}\}.$$

Model the probabilities of these by $\mathrm{P}(K) = p$, $\mathrm{P}(R|K) = 1$ and $\mathrm{P}(R|K^c) = 1/m$. Then apply Bayes rule and the law of total probability to obtain $\mathrm{P}(K^c|R)$:

$$\begin{aligned} \mathrm{P}(K^c|R) &= \frac{\mathrm{P}(R|K^c)\,\mathrm{P}(K^c)}{\underbrace{\mathrm{P}(R)}_{\mathrm{P}(R|K^c)\,\mathrm{P}(K^c)+\mathrm{P}(R|K)\,\mathrm{P}(K)}} \\ &= \frac{(1/m)(1-p)}{(1/m)(1-p)+p} \\ &= 1 - \frac{mp}{1+p(m-1)} \end{aligned}$$

4. If I keep tossing a fair coin, what is the probability I get (a) the pattern HH before the pattern HT; (b) the pattern HH before the pattern TH?

Solution:
(a) Let's say I toss a coin and I haven't come across either patterns HH or HT yet. As soon as I toss an H, I am now interested in the next outcome. Note that whatever happened before this is irrelevant, the only thing that matters is the next coin toss. It is H with probability 1/2 and T with probability 1/2, so the two probabilities are equivalent.

(b) Now I am interested in HH before TH. At any given point, I can toss H or T with equal probability. If I toss a T, then I can never get HH before TH: my next toss is either an H (lose) or another T, and I repeat the argument again. Given that I toss an H, then I toss another H and do so before TH occurs. However if I toss a T then I am in the same situation previously, and I can never get HH before TH again. In other words, out of the four outcomes $\{HH, HT, TH, TT\}$, only 1 is favourable. Therefore, the probability required is 1/4.

5. (a) For independent events $A_1, \ldots, A_n$, show that

$$\mathrm{P}(A_1 \cup \cdots \cup A_n) = 1 - \prod_{i=1}^{n} \big(1 - \mathrm{P}(A_i)\big).$$

Solution: Proof by induction: Clearly $\mathrm{P}(A_1) = 1 - \mathrm{P}(A_1)$. Now consider the probability $\mathrm{P}(A_1 \cup A_2)$. This is equivalent to

$$\begin{aligned}
\mathrm{P}(A_1 \cup A_2) &= \mathrm{P}(A_1) + \mathrm{P}(A_2) - \overbrace{\mathrm{P}(A_1 \cap A_2)}^{\mathrm{P}(A_1)\,\mathrm{P}(A_2)} \\
&= 1 - 1 + \mathrm{P}(A_2) + \mathrm{P}(A_1)\big(1 - \mathrm{P}(A_2)\big) \\
&= 1 - \big(1 - \mathrm{P}(A_1)\big)\big(1 - \mathrm{P}(A_2)\big)
\end{aligned}$$

Now suppose what is needed to be showed is true. Then

$$\begin{aligned}
\mathrm{P}(\overbrace{A_1 \cup \cdots \cup A_n}^{B} \cup A_{n+1}) &= 1 - \big(1 - \mathrm{P}(B)\big)\big(1 - \mathrm{P}(A_{n+1})\big) \\
&= 1 - \left(\prod_{i=1}^{n} \big(1 - \mathrm{P}(A_i)\big)\right)\big(1 - \mathrm{P}(A_{n+1})\big) \\
&= 1 - \prod_{i=1}^{n+1} \big(1 - \mathrm{P}(A_i)\big).
\end{aligned}$$

We have showed that if the hypothesis is true for n, it is true for $n+1$, and we have showed also that it is true for $n = 1$ and $n = 2$. The proof by induction is complete.

(b) A pair of dice is rolled n times. How large must n be so that the probability of rolling at least one double six is more than 1/2?

Solution: Let A_k be the event that you roll a double six on the k-th roll. Note that $\mathrm{P}(A_k) = 1/36$, and assume that each rolls is independent of each other. We are interested in the probability that a double six is rolled at least once in n rolls, i.e. $\mathrm{P}(A_1 \cup \cdots \cup A_n)$. From the above subpart, we have that

$$\mathrm{P}(A_1 \cup \cdots \cup A_n) = 1 - \prod_{i=1}^{n} \big(1 - 1/36\big).$$

We would like this probability to be at least 1/2, so we solve for n from the following:

$$\begin{aligned} 1 - \prod_{i=1}^{n} \big(1 - 1/36\big) &> 1/2 \\ \Rightarrow \big(35/36\big)^n &< 1/2 \\ \Rightarrow n \log(35/36) &< \log(1/2) \\ \Rightarrow n &> \frac{\log(1/2)}{\log(35/36)} = 24.6 \end{aligned}$$

So minimally $n = 25$ for this to occur.

6. Consider a simple random walk on the integers $\{0, 1, ..., 9, 10\}$, with steps ± 1 each with probability $1/2$, and stopped as soon as the walk reaches either 0 or 10. Let T be the number of steps before the walk reaches either 0 or 10. Suppose that $0 \leq a \leq 10$ and set $m(a) = E(T|\text{walk starts at } a)$.

   (a) Explain why $m(0) = m(10) = 0$.

   Solution: When starting at 0 or 10, the walk "stops", so there are no steps to be taken.

   (b) Argue that

   $$m(a) = 1 + \frac{1}{2}m(a-1) + \frac{1}{2}m(a+1) \text{ for } 0 < a < 10.$$

   Solution: When starting at any integer besides 0 and 10, at least one step must be taken to reach 0 or 10 (e.g. $1 \to 0$ or $9 \to 10$). But suppose the next step doesn't end the walk, then one has reached either integer $a+1$ or $a-1$ with equal probability, and the number of steps to end the walk in each case is $m(a+1)$ and $m(a-1)$ respectively. So the equation follows.

   (c) Show that $m(a) = (10 - a)a$ solves these equations.

Solution: Plugging in the solution to the RHS, we get

$$\begin{aligned}\text{RHS} &= 1 + \frac{(10-a+1)(a-1)}{2} + \frac{(10-a-1)(a+1)}{2} \\ &= 1 + \frac{11a - 11 - a^2 + a}{2} + \frac{9a + 9 - a^2 - a}{2} \\ &= 1 + \frac{20a - 2 - 2a^2}{2} \\ &= (10-a)a\end{aligned}$$

(d) Is it the unique solution?

Solution: Realise that we have 9 equations in 9 unknowns, i.e.

$$\begin{aligned}m(1) - m(2)/2 - \cancel{m(0)/2} &= 1 \\ m(2) - m(3)/2 - m(1)/2 &= 1 \\ &\vdots \\ m(9) - \cancel{m(10)/2} - m(8)/2 &= 1\end{aligned}$$

This can be written in matrix form as follows:

$$\underbrace{\begin{pmatrix} 1 & -\frac{1}{2} & & & & \\ -\frac{1}{2} & 1 & -\frac{1}{2} & & & \\ & -\frac{1}{2} & 1 & -\frac{1}{2} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\frac{1}{2} & 1 & -\frac{1}{2} \\ & & & & -\frac{1}{2} & 1 \end{pmatrix}}_{\mathbf{A}\in\mathbb{R}^{9\times 9}} \underbrace{\begin{pmatrix} m(1) \\ \vdots \\ m(9) \end{pmatrix}}_{\mathbf{m}} = \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}}_{\mathbf{1}}$$

The system of linear equations has a unique solution if and only if the determinant of $\mathbf{A}$ is non-zero (i.e. it is invertible). Clearly $\det(\mathbf{A}) = 1 > 0$, so the solution is unique.

7. For a random variable X with mean μ and variance $\mathrm{Var}(X)$ and any given constant $c \in \mathbb{R}$, prove that
   (a) $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mu^2$.
   (b) $\mathrm{Var}(X) = \mathrm{E}\left(X(X-1)\right) + \mu - \mu^2$.
   (c) $\mathrm{E}\left((X-c)^2\right) = \mathrm{Var}(X) + (\mu - c)^2$ so that the minimum mean squared deviation occurs when $c = \mu$.

Solution:
(a)

$$\begin{aligned}\text{Var}(X) &= \text{E}((X-\mu)^2) \\ &= \text{E}(X^2 + \mu^2 - 2\mu X) \\ &= \text{E}(X^2) + \mu^2 - 2\mu \cdot \mu \\ &= \text{E}(X^2) - \mu^2\end{aligned}$$

(b)

$$\begin{aligned}\text{E}\big(X(X-1)\big) + \mu - \mu^2 &= \text{E}(X^2) - \cancel{\text{E}(X) + \mu} - \mu^2 \\ &= \text{Var}(X)\end{aligned}$$

(c)

$$\begin{aligned}\text{E}\big((X-c)^2\big) &= \text{E}(X^2) + c^2 - 2c\mu \\ &= \text{E}(X^2) + c^2 - 2c\mu - \mu^2 + \mu^2 \\ &= \text{Var}(X) + (c-\mu)^2\end{aligned}$$

8. (a) By example, or otherwise, show that generally $\text{E}[\phi(X)] \neq \phi(\text{E}[X])$.

Solution: A simple example is the following:

$$X = \begin{cases} 1 & \text{w.p. } 1/2 \\ 3 & \text{w.p. } 1/2 \end{cases} \qquad \Rightarrow \qquad 1/X = \begin{cases} 1 & \text{w.p. } 1/2 \\ 1/3 & \text{w.p. } 1/2 \end{cases}$$

and $\text{E}(X) = 2/3 \neq 1/2 = \text{E}(1/X)$.

(b) A game is presented to you as follows: Independent random variables X_i whose values generated by a computer take on either 0.5 with probability $\frac{1}{i+1}$, or 1 otherwise, for $i = 1, \dots, 5$. These values are then multiplied together to give $X = X_1 X_2 \cdots X_5$, and $Y = 1/X$ is calculated. You are returned B\$ Y for playing this game, after paying a certain fee to play. What is the maximum fee you are willing to pay to play this game?

Solution: The key is finding out what is the average pay out of this game, $\text{E}(Y)$. Let

$$Y_i = 1/X_i = \begin{cases} 1 & \text{w.p. } \frac{i}{i+1} \\ 2 & \text{w.p. } \frac{1}{i+1} \end{cases}$$

Since each Y_i is independent and $\mathrm{E}(Y_i) = \frac{i+2}{i+1}$, we have that

$$\begin{aligned}
\mathrm{E}(Y) &= \mathrm{E}\Big(\prod_{i=1}^{5} Y_i\Big) \\
&= \prod_{i=1}^{5} \mathrm{E}(Y_i) \\
&= 7/2
\end{aligned}$$

So the expected return is B\$3.50. You shouldn't be willing to pay any more than this to play this game.

9. The number of insurance claims that will be made directly to your company in each of n counties next month are modelled as n independent random variables $X_i \sim \mathrm{Pois}(\theta_i)$, $i = 1, \dots, n$. Write $\psi = \sum_{i=1}^{n} \theta_i$. The total monthly direct claims is modelled as the random variable $X = \sum_{i=1}^{n} X_i$.

   (a) Obtain the probability generating function of X_i, and hence of X. Deduce the distribution of X.

   Solution:

$$\begin{aligned}
G_{X_i}(s) &= \mathrm{E}(s^X) \\
&= \sum_{x=0}^{\infty} s^x \cdot \frac{e^{-\theta_i}\theta_i^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-s\theta_i}(s\theta_i)^x}{x!} \cdot e^{s\theta_i} e^{-\theta_i} \\
&= e^{\theta_i(s-1)}
\end{aligned}$$

   Therefore, the PGF of $X = X_1 + \cdots + X_n$ is $G_X(s) = \prod_{i=1}^{n} G_{X_i}(s) = \exp\{\sum_{i=1}^{n} \theta_i(s-1)\} = \exp\{\psi(s-1)\}$. This shows that $X \sim \mathrm{Pois}(\psi)$ since PGF uniquely characterises the distribution.

   (b) The number of indirect claims for next month is modelled as an independent random variable W, with PGF $G_W(s) = e^{\psi(s^2-1)}$. Obtain the PGF of the total claims $Y = X + W$, $\mathrm{E}(Y)$, and $\mathrm{Var}(Y)$.

   Solution: The PGF of Y is $G_Y(s) = G_X(s)G_W(s) = \exp\{\psi(s-1)\}\exp\{\psi(s^2-1)\} = \exp\{\psi(s^2+s-2)\}$. So $\mathrm{E}(Y) = G'_Y(s)|_{s=1} = \psi(2s+1)G_Y(s)|_{s=1} = 3\psi$. We also have the result stating $\mathrm{E}(Y(Y-1)) = G''_Y(s)|_{s=1}$. Since $G''_Y(s)|_{s=1} = (2\psi+\psi^2(2s+1)^2)G_Y(s)|_{s=1} = 2\psi+9\psi^2$, $\mathrm{Var}(Y) = 2\psi+9\psi^2+3\psi-(3\psi)^2 = 5\psi$.

10. Conditional upon an unknown scientific constant μ, let $X_i \sim f$ be iid random variables representing the future outcomes of a series of experiments, with $\mathrm{E}(X_i) = \mu$ units and $\mathrm{Var}(X_i) = 400$ squared units. The estimator for μ will be $\bar{X} = \sum_{i=1}^n X_i/n$. Assuming that the CLT applied with sufficiently fast convergence,

(a) What is the probability that the realisation of $\bar{X}$ will be within 1 unit of μ when

i. $n = 1$?

ii. $n = 4$?

iii. $n = 16$?

iv. $n = 100$?

Solution: Using the CLT, we have that $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i \approx \mathrm{N}(\mu, 400/n)$, so the probability of interest is

$$\mathrm{P}(|\bar{X}_n - \mu| < 1) = \mathrm{P}\left(\overbrace{\left| \frac{\bar{X}_n - \mu}{\sqrt{400/n}} \right|}^{\sim \mathrm{N}(0,1)} < \frac{1}{\sqrt{400/n}} \right) = 2\Phi(\sqrt{n}/20) - 1$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. By looking up the tables, we find that the probabilities are as follows: i. 0.0399, ii. 0.0797, iii. 0.159, iv. 0.383.

(b) If the experimenter asks you what is the least number of experiments that should be performed in order to have a probability of at least 0.95 that $\bar{X}$ will be within 2 units of μ, what do you reply?

Solution: This requires solving for n in the equation below:

$$\begin{aligned} 2\Phi(2\sqrt{n}/20) - 1 &\geq 0.95 \\ \Rightarrow \Phi(\sqrt{n}/10) &\geq 1.95/2 = 0.975 \\ \Rightarrow \sqrt{n}/10 &\geq \Phi^{-1}(0.975) = 1.96 \\ \Rightarrow n &\geq 384.16 \end{aligned}$$

So minimally, 385 experiments should be run.