1. Jane, an avid astronomer, invites her friend Phyllis, a statistician, to observe a meteor shower. Phyllis, indifferent to astronomy but curious about the statistical aspects, inquires further. She learns that during this meteor shower Jane claims that meteors will occur at an average rate of one meteor every three minutes. Furthermore at most one meteor appears at a given moment and each meteor's appearance is independent of the appearance of other meterors.

   Intrigued by the opportunity to test this statistical claim, Phyllis decides to come along. She decides to carry out the test by counting the number of meteors in one hour.

   (a) Let $X$ represent the total number of meteors observed in that hour. Given Jane's claim of an average rate of one meteor every three minutes, define the distribution of $X$ and its parameters. Please also state any additional assumptions you have made about the situation. [2 marks]

   (b) Given Jane's assumptions about average rate, state the null and alternative hypotheses concerning the distribution of $X$. Specify the full distribution of $X$ and use a two-sided alternative hypothesis. [2 marks]

   (c) Phyllis counts 12 meteors in the hour of observation. Sketch as a curve the shape of the probability function of $X$ under the null hypothesis. Your sketch should have axes labelled $x$ and $\mathbb{P}(X = x)$. Indicate on the sketch the upper and lower limits of $x$, and the approximate value of $x$ where the curve peaks under the null hypothesis. Also mark the observed value $x$ so that you can see the tail probabilities required for the p-value, and shade under the curve the area represented by the p-value. [4 marks]

   (d) Describe a method suitable for determining the $p$-value and implement this technique. Interpret the result in terms of the strength of evidence against the null hypothesis. From your analysis, can Phyllis confidently challenge Jane's assertion about the average number of meteors observed within an hour?
   ***Hint:*** You might consider using the `R` software. If so, write down the appropriate `R`-command, run it in `R`, and interpret the outcome. [4 marks]

[Total: 12 marks]

2. Let $X$ and $Y$ be two independent random variables with the property that $X \sim \text{Bernoulli}(p_1)$ and $Y \sim \text{Bernoulli}(p_2)$, where $0 < p_1, p_2 < 1$.

   We define $U = \max(X, Y)$ and $V = \min(X, Y)$.

   As $X$ and $Y$ can only take values 0 and 1, we note that:
   $\max(0, 0) = 0$,
   $\max(0, 1) = \max(1, 0) = \max(1, 1) = 1$,
   $\min(0, 0) = \min(0, 1) = \min(1, 0) = 0$ and
   $\min(1, 1) = 1$.

   Answer the following questions:

   (a) Find the cdf of $U$. The distribution of $U$ is one of the named distributions described in the course book. State which one, including parameter/parameters.
   Hint: Copy the following chain of identities to your answer and replace $\ldots$ with the correct justifications. Use these identities for finding the cdf of $U$.
   For $u \in \mathbb{R}$, we have:

$$\begin{aligned} F_U(u) &= \mathbb{P}(U \leq u) \; [\ldots] \\ &= \mathbb{P}(\max(X, Y) \leq u) \; [\ldots] \\ &= \mathbb{P}(X \leq u, Y \leq u) \; [\ldots] \\ &= \mathbb{P}(X \leq u)\mathbb{P}(Y \leq u) \; [\ldots] \end{aligned}$$

   [6 marks]

   (b) Find the cdf of $V$. The distribution of $V$ is one of the named distributions described in the course book. State which one, including parameter/parameters.
   Hint: Copy the following chain of identities to your answer and replace $\ldots$ with the correct justifications. Use these identities for finding the cdf of $V$.
   For $v \in \mathbb{R}$, we have:

$$\begin{aligned} F_V(v) &= \mathbb{P}(V \leq v) \; [\ldots] \\ &= \mathbb{P}(\min(X, Y) \leq v) \; [\ldots] \\ &= 1 - \mathbb{P}(\min(X, Y) > v) \; [\ldots] \\ &= 1 - \mathbb{P}(X > v, Y > v) \; [\ldots] \\ &= 1 - \mathbb{P}(X > v)\mathbb{P}(Y > v) \; [\ldots] \\ &= 1 - [1 - \mathbb{P}(X \leq v)][1 - \mathbb{P}(Y \leq v)] \; [\ldots] \end{aligned}$$

   [6 marks]

(c) The joint probability mass function of a set of random variables is the function that describes every possible set of outcomes. Find the joint probability mass function (pmf) for the random variables $U$ and $V$. In other words, find the probability of every possible set of outcomes $(U, V) = (u, v)$ Show all your calculations.

<u>Hint</u>: Remember that, for example,

$$\mathbb{P}(U = 0, V = 1) = \mathbb{P}((U, V) = (0, 1))$$

$$= \mathbb{P}(\max(X, Y) = 0, \min(X, Y) = 1),$$

and as it is impossible to have $\max(X, Y) < \min(X, Y)$, it follows that $\mathbb{P}(U = 0, V = 1) = 0$.

[6 marks]

(d) Decide whether the random variables $U$ and $V$ are independent.

<u>Hint</u>: Apply the definition of independence from Section 1.2.4 of the course book.

[2 marks]

[Total: 20 marks]

3. In each of the following cases the random variable $X$ is one of the named distributions described in Table 2.1 in the coursebook. In each case state the distribution of $X$, including parameters. Show your workings, and clearly state any additional assumptions that you have made.

(a) Jennifer loves to solve challenging math problems. Every day, she tries to solve a number of them, and she has a success rate of 0.7 for any given problem. Let $X$ be the number of problems she has to attempt before she successfully solves 5 problems. [3 marks]

(b) Paul is trying to capture a rare photo of a specific bird. Each day he goes out, there's a 0.2 probability that he'll capture a good shot of the bird. Let $X$ be the number of days Paul tries to get the shot before the day he finally get his first good shot [3 marks]

(c) In a standard deck of 52 playing cards, there are 4 aces. Samantha randomly selects 5 cards without replacement. Let $X$ be the number of aces in her selection. [3 marks]

(d) At a factory, a machine produces light bulbs. Historically, 5% of the bulbs produced by this machine are defective. If 20 bulbs are randomly selected from a day's production, let $X$ be the number of defective bulbs. What is the distribution of $X$? [3 marks]

[Total: 12 marks]

4. Charles, a dog socialization trainer, always works with dogs in pairs made up of one male and one female. Charles' method involves teaching the dogs a trick of their choosing. The dogs pick solely based on predisposition, and learns either bowing or twirling. Charles has found that each dog is equally likely to pick either trick.

   Occasionally, with probability $\alpha$, end up becoming really good friends, and end up not only performing the same trick, but also perform it at the the same time, in perfect synchrony.

   **Notation:**

| | |
|---|---|
| B | Dog chooses bowing. |
| T | Dog chooses twirling. |
| Syn | Pair performs in synchrony. |

   For documentation, Charles notes the female dog's trick first. Thus, TB means the female twirls and the male bows, while BT indicates the opposite. BB and TT denote both dogs bowing or twirling, respectively (but does not necessarily indicate that the dogs perform in synchrony).

   (a) State $\mathbb{P}(B)$ and $\mathbb{P}(T)$ [1 marks]

   (b) Show that the probability that the two dogs in a pair do different tricks is $\dfrac{1-\pi}{2}$.
   ***Hint:*** Be very careful with the notation, in particular with identifying conditional and joint events. Also, notice that dog pairs who are not in sync have four possible outcomes, while dogs who are in sync have a different number of possible outcomes. [4 marks]

   (c) Explain why it seems reasonable that the dogs who are in sync are as likely to both bow as they are to both twirl. Find this probability. [2 marks]

   (d) Find the probabilities $\mathbb{P}(TT)$ and $\mathbb{P}(BB)$. [4 marks]

[Total: 11 marks]

**Total Homework 1: 55 marks.**