MH1820 Introduction to Probability and Statistical Methods

Tutorial 2 (Week 3) Solution

Problem 1 (sample spaces)

For each of the following statistical experiments, write down a suitable sample space.

(i) The number of passengers on an MRT train is counted.

(ii) The height of a person is measured.

(iii) 1000 persons are selected randomly and they are asked if they have any travel plans for this year.

(iv) The maximum force (in N=Newton) is determined that a certain steel beam can withstand without breaking.

(v) At a random time, the number of seconds past the current minute is determined (for instance, if the current time is 12:13:33, the outcome is 33).

(vi) A dice is rolled 5 times.

Solution Sample spaces are not uniquely determined, so the solutions given here are not the only valid ones. Typically, it is best to keep the sample space as simple as possible while making sure that it represents all possible outcomes of the experiment.

(i) Here the simplest choice for the sample space is $\Omega = \mathbb{Z}_0^+ = \{0, 1, 2, \ldots\}$ (set of nonnegative integers). An outcome $x \in \Omega$ means that there are x passengers on the train.

Since usually there are less than 2000 passengers on an MRT train, we could also choose $\{0, 1, \ldots, 2000\}$ as the sample space, but an upper bound like 2000 is not useful for calculating probabilities. When large numbers of possible outcomes are involved, we typically only use *approximations* of probabilities (not exact probabilities) and for these approximations we do not need an upper bound.

(ii) The height of a person can be represented by a nonnegative real number (the height in centimeters), so we can take $\Omega = \mathbb{R}_0^+$ (set of nonnegative real numbers) as the sample space. Again, we could introduce an upper bound (say $\Omega = [0, 300]$, the interval from 0 to 300), but this is not necessary.

(iii) Let $P_1, \ldots, P_{1000}$ be the selected persons and set $x_i = 0$ if P_i answers no and $x_i = 1$ if P_i answers yes. Given this setting, Ω can be taken as the set of all 1000-tuples $(x_1, \ldots, x_{1000})$ with $x_i \in \{0, 1\}$. Note that this sample space is huge ($|\Omega| = 2^{1000}$).

(iv) Given no further information, the best choice for the sample space is $\Omega = \mathbb{R}_0^+$.

(v) The possible outcomes are $0, 1, \dots, 59$ and we can take $\Omega = \{0, 1, \dots, 59\}$. If the measurement was more precise, say up to milliseconds, then $\Omega = [0, 60]$ (the interval of real numbers from 0 to 60) would be more appropriate.

(vi) Let $x_i \in \{1, \dots, 6\}$ be the outcome of the ith roll. In this setting, the sample space Ω is the set of all 5-tuples $(x_1, \dots, x_5)$ with $x_i \in \{1, \dots, 6\}$. Note that $|\Omega| = 6^5$.

Problem 2 (independent events)

Two integers are drawn independently at random from $1, 2, \dots, 100$ (repetition is allowed, for instance, both numbers could be 1). Consider the following events.

E_1: both numbers are even
E_2: the first number is ≤ 50
E_3: the second number is prime

For any pair of events above, decide whether the two events are independent.

$\boxed{\textbf{Solution}}$ To determine if two events A, B are independent, we have to check if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Hence we need to compute all probabilities $\mathbb{P}(E_i)$ and $\mathbb{P}(E_i \cap E_j)$. The sample space Ω is the set of 2-tuples (x_1, x_2) with $x_1, x_2 \in \{1, \dots, 100\}$ and we have $|\Omega| = 100^2 = 10000$.

- $\mathbb{P}(E_1)$: There are 50 even numbers in $\{1, \dots, 100\}$. By the multiplication principle, $|E_1| = 50^2$ and thus $\mathbb{P}(E_1) = 50^2/10000 = 0.25$.
- $\mathbb{P}(E_2)$: By the multiplication principle, $|E_2| = 50 \cdot 100$ and thus $\mathbb{P}(E_2) = 5000/10000 = 0.5$.
- $\mathbb{P}(E_3)$: There are 25 primes in $\{1, \dots, 100\}$. By the multiplication principle, $|E_3| = 100 \cdot 25$ and thus $\mathbb{P}(E_3) = 2500/10000 = 0.25$.
- $\mathbb{P}(E_1 \cap E_2)$: $E_1 \cap E_2$ is the event that x_1 is even and ≤ 50 and x_2 is even. There are 25 even numbers ≤ 50 and 50 even numbers ≤ 100. Hence $|E_1 \cap E_2| = 25 \cdot 50 = 1250$ and $\mathbb{P}(E_1 \cap E_2) = 1250/10000 = 0.125$.
- $\mathbb{P}(E_1 \cap E_3)$: $E_1 \cap E_3$ is the event that both x_1 and x_2 are even and that x_2 is a prime. Since 2 is the only even prime, there is only one choice for x_2. Since there are 50 choices for x_1, we have $|E_1 \cap E_3| = 50$ and $\mathbb{P}(E_1 \cap E_3) = 0.005$.
- $\mathbb{P}(E_2 \cap E_3)$: Note that $E_2 \cap E_3$ is the event that $x_1 \leq 50$ and x_2 is a prime. There are 50 choices for x_1 and 25 for x_2. Hence $|E_2 \cap E_3| = 50 \cdot 25 = 1250$ and $\mathbb{P}(E_2 \cap E_3) = 0.125$.

In summary we have

$$
\begin{array}{ll}
\mathbb{P}(E_1 \cap E_2) = 0.125 & \mathbb{P}(E_1)\mathbb{P}(E_2) = 0.25 \cdot 0.5 = 0.125 \\
\mathbb{P}(E_1 \cap E_3) = 0.005 & \mathbb{P}(E_1)\mathbb{P}(E_3) = 0.25 \cdot 0.25 = 0.0625 \\
\mathbb{P}(E_2 \cap E_3) = 0.125 & \mathbb{P}(E_2)\mathbb{P}(E_3) = 0.5 \cdot 0.25 = 0.125
\end{array}
$$

Hence the only pairs of independent events are $\{E_1, E_2\}$ and $\{E_2, E_3\}$.

$\square$

Problem 3 (conditional probability)

A researcher finds that, of 982 men who died in 2002, 221 died from some heart disease. Also, of the 982 men, 334 had at least one parent who had some heart disease. Of the latter 334 men, 111 dies from some heart disease. A man is selected from the group of 982. Given that neither of his parents had some heart disease, find the conditional probability that this man died of some heart disease.

$\boxed{\textbf{Solution}}$ Let A be the event that a man died from heart disease, and B be the event that at least one parent of the man had some heart disease. The following table gives the size of the events and their respective intersections from the information given.

	A	$\overline{A}$	TOTAL
B	111	223	334
$\overline{B}$	110	538	648
TOTAL	221	761	982

We want to compute $\mathbb{P}(A|\overline{B})$. Hence,

$$\mathbb{P}(A|\overline{B}) = \frac{\mathbb{P}(A \cap \overline{B})}{\mathbb{P}(\overline{B})} = \frac{\frac{110}{982}}{\frac{648}{982}} = \frac{110}{648}.$$

$\square$

Problem 4 (conditional probability)

(a) A dice is rolled 4 times.

(i) What is the probability that **at most** one of the rolls is a 6?

(ii) What is the probability that **at least** two of the rolls are a 6?

(iii) What is the probability that the total of the rolls is **at least** 22?

(iv) What is the probability that the total of the rolls is **at least** 22 under the condition that at least two of the rolls are a 6?

(b) A family has two children. Assume that the probability for a child to be a girl is $\frac{1}{2}$, and the probability for a child to be a boy is $\frac{1}{2}$.

(i) What is the probability that both children are girls under the condition that the first child is a girl?

(ii) What is the probability that both children are girls under the condition at least one of the children is a girl?

$\boxed{\textbf{Solution}}$ (a) We take the set Ω of all 4-tuples (x_1, x_2, x_3, x_4) with $x_i \in \{1, \dots, 6\}$ as sample space. Note that $|\Omega| = 6^4 = 1296$.
Let B be the event that at least two of the rolls are a 6. Then $\overline{B}$ is the event that at most one of the rolls is a 6.

(i) The number of outcomes without any 6 is 5^4, since there are 5 possibilities different from 6 for each roll. The number of outcomes with exactly one 6 is $\binom{4}{1} \cdot 5^3$ (there are $\binom{4}{1}$ possibilities to choose the positions of the 6 and 5^3 possibilities to fill in the remaining three positions). Hence $|\overline{B}| = 5^4 + 4 \cdot 5^3 = 1125$ and $\mathbb{P}(\overline{B}) = 1125/1296$.

(ii) By the properties of probability,

$$\mathbb{P}(B) = 1 - \mathbb{P}(\overline{B}) = 1 - 1125/1296 = 171/1296.$$

(iii) To get a total of at least 22, there are the following possibilities.

- Exactly two 6's and exactly two 5's. There are $\binom{4}{2} = 6$ choices for the positions of the 6's among (x_1, x_2, x_3, x_4). Hence there are exactly 6 such outcomes with exactly two 6's and exactly two 5's.
- Exactly three 6's together with a 4 or a 5. There are $\binom{4}{3}$ possibilities to choose the positions of the 6's and two choices for the remaining number (4 or 5). Hence there are exactly $\binom{4}{3} \cdot 2 = 8$ such outcomes with exactly three 6's.
- Exactly four 6's. There is only one outcome like this.

Let A be the event to roll a total of at least 22. Then

$$\mathbb{P}(A) = \frac{6 + 8 + 1}{1296} = \frac{15}{1296}.$$

(iv) Notice that $A \subseteq B$. So $\mathbb{P}(A \cap B) = \mathbb{P}(A) = \frac{15}{1296}$. We conclude that

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{15}{1296}}{\frac{171}{1296}} = \frac{15}{171} \approx 8.8\%.$$

Remark: When computing $|B|$, there is a danger of erroneous double counting if the following approach is used. There are $\binom{4}{2}$ positions for two 6's and 6^2 ways to fill in the remaining two positions (since the condition is "*at least*" two 6's, there still can be 6's in the remaining positions). This way of counting would give the wrong answer $|B| = 6 \cdot 6^2 = 216$. For instance, the outcome $(6, 6, 6, 6)$ would have been counted 6 times with this approach.

(b) Let $x_i = G$ if the ith child is a girl and $x_i = B$ if the ith child is a boy. The sample space is $\Omega = \{(G, G), (G, B), (B, G), (B, B)\}$ and each outcome (x_1, x_2) has probability $1/4$. The relevant events are

A: both children are girls
B_1: the first child is a girl
B_2: at least one of the children is a girl

Note that

$$\begin{aligned} A &= \{(G,G)\}, \\ B_1 &= \{(G,G),(G,B)\}, \\ B_2 &= \{(G,G),(G,B),(B,G)\}. \end{aligned}$$

Hence $\mathbb{P}(A \cap B_1) = \mathbb{P}(A \cap B_2) = \mathbb{P}(A) = |A|/|\Omega| = 1/4$, $\mathbb{P}(B_1) = 2/4 = 1/2$, and $\mathbb{P}(B_2) = 3/4$. We conclude

$$\begin{aligned} \mathbb{P}(A|B_1) &= \frac{\mathbb{P}(A \cap B_1)}{\mathbb{P}(B_1)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}, \\ \mathbb{P}(A|B_2) &= \frac{\mathbb{P}(A \cap B_2)}{\mathbb{P}(B_2)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}. \end{aligned}$$

Thus

(i) the probability that both children are girls under the condition that the first child is a girl is 1/2;

(ii) the probability that both children are girls under the condition at least one of the children is a girl is 1/3.

□

Problem 5 (law of total probability and Bayes' theorem)

A ball is drawn from one of 3 boxes. The boxes contain the following number of balls of colors blue (B), red (R), and yellow (Y).

	B	R	Y
Box 1	1	4	5
Box 2	6	2	2
Box 3	3	3	4

The following procedure is used to draw the ball.

- One of the boxes is chosen at random: Box 1 is chosen with probability 0.2, Box 2 is chosen with probability 0.5, and Box 3 is chosen with probability 0.3.
- A ball is drawn from the chosen box (each ball in the box is chosen with the same probability).

(i) What is the probability that a blue ball is drawn?

(ii) What is the probability that a yellow ball is drawn?

(iii) If a blue ball is drawn, what is the probability that it was drawn from Box 1?

$\boxed{\textbf{Solution}}$ We first define the relevant events.

$$\begin{aligned} B_1, B_2, B_3 &: \text{ball is drawn from box } 1, 2, 3, \text{ respectively,} \\ A &: \text{a blue ball is drawn,} \\ Y &: \text{a yellow ball is drawn.} \end{aligned}$$

Since the probability that the ball is drawn from box 1 is 0.2, we have $\mathbb{P}(B_1) = 0.2$. Similarly, $\mathbb{P}(B_2) = 0.5$ and $\mathbb{P}(B_3) = 0.3$.

If the ball is drawn from box 1, then the probability that it is blue is 0.1, since 1 out of 10 balls in box 1 is blue. Hence $\mathbb{P}(A|B_1) = 0.1$. Similarly, we get

$$\begin{aligned} \mathbb{P}(A|B_1) &= 0.1, \\ \mathbb{P}(A|B_2) &= 0.6, \\ \mathbb{P}(A|B_3) &= 0.3, \\ \mathbb{P}(Y|B_1) &= 0.5, \\ \mathbb{P}(Y|B_2) &= 0.2, \\ \mathbb{P}(Y|B_3) &= 0.4. \end{aligned}$$

(i) By the Law of Total Probability, the probability that a blue ball is drawn is

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) + \mathbb{P}(A|B_3)\mathbb{P}(B_3) = 0.1 \cdot 0.2 + 0.6 \cdot 0.5 + 0.3 \cdot 0.3 = 0.41.$$

(ii) By the Law of Total Probability, the probability that a yellow ball is drawn is

$$\mathbb{P}(Y) = \mathbb{P}(Y|B_1)\mathbb{P}(B_1) + \mathbb{P}(Y|B_2)\mathbb{P}(B_2) + \mathbb{P}(Y|B_3)\mathbb{P}(B_3) = 0.5 \cdot 0.2 + 0.2 \cdot 0.5 + 0.4 \cdot 0.3 = 0.32.$$

(iii) By Bayes' Theorem, under the condition that a blue ball was drawn, the probability that it was drawn from box 1 is

$$\mathbb{P}(B_1|A) = \frac{\mathbb{P}(A|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A)} = \frac{0.1 \cdot 0.2}{0.41} \approx 0.049.$$

$\square$

Answer Keys. 2. $\mathbb{P}(E_1) = 0.25$, $\mathbb{P}(E_2) = 0.5$, $\mathbb{P}(E_3) = 0.25$, $\mathbb{P}(E_1 \cap E_2) = 0.125$, $\mathbb{P}(E_1 \cap E_3) = 0.005$, $\mathbb{P}(E_2 \cap E_3) = 0.125$. 3. $\frac{110}{648}$ 4. (a) (i) $\frac{1125}{1296}$ (ii) $\frac{171}{1296}$ (iii) $\frac{15}{1296}$ (iv) $\frac{15}{171}$ (b) (i) $\frac{1}{2}$, (ii) $\frac{1}{3}$. 5. (i) 0.41 (ii) 0.32 (iii) 0.049.