PROBABILITY PRACTICAL 1 SOLUTIONS

DAVID STEINSALTZ

(1) Blood groups are distributed in the UK as follows:

O	A	B	AB
48%	39%	10%	3%

(a) If two people are picked at random from the population, what is the chance that their blood is of the same type? Of different types?

$$\begin{aligned}\mathbb{P}\{\text{same type}\} &= \mathbb{P}\{\text{both A}\} + \mathbb{P}\{\text{both B}\} + \mathbb{P}\{\text{both O}\} + \mathbb{P}\{\text{both AB}\}\\ &= 0.48^2 + 0.39^2 + 0.10^2 + 0.03^2\\ &= 0.39.\end{aligned}$$

Also

$$\mathbb{P}\{\text{different type}\} = 1 - \mathbb{P}\{\text{same type}\} = 0.61.$$

(b) If four people are picked at random, let p_k be the probability that there are exactly k different blood types among them. Find all values of p_k.

$$\begin{aligned}p_1 &= \mathbb{P}\{\text{all A}\} + \mathbb{P}\{\text{all B}\} + \mathbb{P}\{\text{all O}\} + \mathbb{P}\{\text{all AB}\}\\ &= 0.48^4 + 0.39^4 + 0.10^4 + 0.03^4\\ &= 0.076.\\ p_4 &= 4! \cdot 0.48 \cdot 0.39 \cdot 0.10 \cdot 0.03\\ &= 0.013.\\ p_3 &= 12\Big(0.48 \cdot 0.39 \cdot 0.10\,(0.48 + 0.39 + 0.10)\\ &\qquad 0.48 \cdot 0.10 \cdot 0.03\,(0.48 + 0.10 + 0.03)\\ &\qquad 0.48 \cdot 0.39 \cdot 0.03\,(0.48 + 0.39 + 0.03)\\ &\qquad 0.03 \cdot 0.39 \cdot 0.10\,(0.03 + 0.39 + 0.10)\Big)\\ &= 0.296.\\ p_2 &= 1 - p_1 - p_3 - p_4\\ &= 0.615.\end{aligned}$$

(2) In a genetic experiment, the offspring of a particular cross have 25% chance of being yellow, and 75% chance of being green.

(a) If there are 10 offspring, calculate the probability that at least 8 are green.

The probability of any one being green is 0.75. So the number of greens has binomial distribution with parameters $(10, 0.75)$. The probability of at least 8 is

$$\sum_{i=8}^{10} \binom{10}{i} (0.75)^i (0.25)^{10-i} = 0.526.$$

(b) Suppose they also have a 25% chance of being short and 75% chance of being tall. Calculate the expected number that are tall and yellow. What assumptions do you need to make?
Assuming the two characters are independent, the probability of being tall and yellow is $0.75 \cdot 0.25 = 0.1875$. So the expected number out of ten trials is 1.875.

(3) Continuing an example from the lecture, suppose there is a disease that occurs in three forms: Mild, Severe, and Lethal. There is a gene that is known to occur in two *alleles* (variants), denoted A_1 and A_2, where the A_1 allele provides some protection against the symptoms of the disease, but does not prevent the disease. 75% of the general population has A_1, and 75% of those with the disease has A_1. Of those people with A_1 who have the disease, 90% have the Mild form, and the rest have the Severe form. The A_2 sufferers are evenly split between the Severe and Lethal forms.

Suppose you observe a patient with the Severe form. Calculate the probability that the patient is of type A_1. Do this with Bayes' Rule and with natural frequencies.

$$\begin{aligned}\mathbb{P}\left\{ \text{ Severe } \middle| A_1 \right\} &= 0.1,\\ \mathbb{P}\left\{ \text{ Severe } \middle| A_2 \right\} &= 0.5,\\ \mathbb{P}(A_1) &= 0.75.\end{aligned}$$

By Bayes Rule,

$$\begin{aligned}\mathbb{P}\{A_1 \mid \text{ Severe } \} &= \frac{\mathbb{P}\left\{ \text{ Severe } \middle| A_1 \right\}\mathbb{P}(A_1)}{\mathbb{P}\left\{ \text{ Severe } \middle| A_1 \right\}\mathbb{P}(A_1) + \mathbb{P}\left\{ \text{ Severe } \middle| A_2 \right\}\mathbb{P}(A_2)}\\ &= \frac{0.1 \times 0.75}{0.1 \times 0.75 + 0.5 \times 0.25}\\ &= 0.375.\end{aligned}$$

Using Natural Frequencies, we consider 1000 randomly chosen patients. 750 are type A_1, and of them 75 have Severe disease. 250 are type A_2, and of them $250 \times 0.5 = 125$ have Severe disease. So there are a total of 200 with Severe disease, of whom 75 are of type A_1, yielding a proportion $75/200 = 0.375$.

(4) I roll a fair die, and then flip a number of fair coins equal to the number that came up on the die.

(a) Calculate the probability that exactly four heads come up on the coins.
Let X be the outcome of the die roll, and Y the number of heads in the coin flips. Then

$$\begin{aligned}\mathbb{P}\{Y = 4\} &= \sum_{i=1}^{6} \mathbb{P}\{X = i\}\mathbb{P}\{Y = 4 \mid X = i\}\\ &= \sum_{i=4}^{6} \frac{1}{6}\binom{i}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^{i-4}\\ &= \frac{1}{6}\left(\frac{1}{16} + \frac{5}{32} + \frac{15}{64}\right)\\ &= \frac{29}{384}\\ &= 0.0755.\end{aligned}$$

(b) Given that three heads come up, calculate the probability that the die roll was 5.

$$\begin{aligned}
\mathbb{P}\{X = 5 \,|\, Y = 4\} &= \frac{\mathbb{P}\{X = 5 \,\&\, Y = 4\}}{\mathbb{P}\{Y = 4\}} \\
&= \frac{5/32 \cdot 1/6}{29/384} \\
&= \frac{10}{29} \\
&= 0.345.
\end{aligned}$$

(5) Suppose we have a sequence of independent trials, each with probability p of success. Let X be the number of the trial on which you have the first success.

(a) Calculate the probability mass function $\mathbb{P}\{X = k\}$.
The event $\{X = k\}$ occurs exactly when there are $k-1$ failures followed by a success. The probability is $\mathbb{P}\{X = k\} = (1-p)^{k-1}p$.

(b) Calculate the expectation and variance of X.
Starting from the formula for the sum of a geometric series

$$\frac{1}{p} = \sum_{k=0}^{\infty} (1-p)^k$$

we take two derivatives to get

$$\begin{aligned}
\frac{1}{p^2} &= \sum_{k=1}^{\infty} k(1-p)^{k-1} \\
\frac{2}{p^3} &= \sum_{k=1}^{\infty} k(k-1)(1-p)^{k-2} = \sum_{k=1}^{\infty} (k+1)k(1-p)^{k-1}.
\end{aligned}$$

And subtracting these we get

$$\frac{2}{p^3} - \frac{1}{p^2} = \sum_{k=1}^{\infty} k^2(1-p)^{k-1}.$$

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k\mathbb{P}\{X = k\} \\
&= \sum_{k=1}^{\infty} kp(1-p)^{k-1} \\
&= p \cdot \frac{1}{p^2} \\
&= \frac{1}{p}.
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=1}^{\infty} k^2 \mathbb{P}\{X = k\} \\
&= p\left(\frac{2}{p^3} - \frac{1}{p^2}\right) \\
&= \frac{2}{p^2} - \frac{1}{p} \\
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{1}{p^2} - \frac{1}{p} \\
&= \frac{1-p}{p^2}
\end{aligned}
$$