Problem Sheet 2

Please send any comments or corrections to `martin@stats.ox.ac.uk`.

1. Two dice are thrown and the numbers they show are represented by a pair (i, j). Suppose that the 36 possible outcomes in $\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$ are all equally likely. Let $A = \{\text{the first die shows 3}\}$, $B = \{\text{the second die shows an even number}\}$ and $C = \{\text{the sum is even}\}$.
   (a) Show that A and B are independent.
   (b) Show that B and C are independent.
   (c) Show that A and C are independent.
   (d) Are A, B and C independent?

2. An urn contains m red balls and n blue balls. Two balls are drawn uniformly at random from the urn, without replacement.
   (a) What is the probability that the first ball drawn is red?
   (b) What is the probability that the second ball drawn is red?
   (c) What is the probability that the first is red given that the second is red?

3. Parliament contains a proportion p of Conservative members, who never update their views about anything, and a proportion $1 - p$ of Labour members, who change their minds at random (with probability r) between successive votes on the same issue. A randomly chosen member is noticed to have voted twice in succession in the same way. What is the probability that this member will vote the same way next time?

4. (a) A fair coin is tossed 26 times. Write down an expression for the probability of seeing exactly 13 heads and 13 tails.
   (b) A pack of 52 cards (containing 26 red and 26 black cards) is shuffled, and then 26 cards are dealt. Write down an expression for the probability that exactly 13 red and 13 black cards are dealt.

   Without calculating, which of the answers in (a) and (b) do you think will be bigger?

   Stirling's formula gives $n! \approx \sqrt{2\pi}\, e^{-n}\, n^{n+1/2}$. Check your intuition by using this formula to approximate the two probabilities.

5. Suppose we have the following rates:
   - a prevalence of Covid infection in the population of 1.25%;
   - for lateral flow tests (LFTs), a false positive rate of 0.07% and a false negative rate of 45%;
   - for PCR tests, a false positive rate of 0.05% and a false negative rate of 7%.

   (Of course it's hard to pin down such rates with any confidence. Most notably, it's not particularly appropriate to give a single number for the false negative rates – the chance of a positive result for an infected person is very much influenced by the time since their infection, for example, and by numerous other factors. But we will work with these figures for the purpose of this question.)

(a) A randomly chosen individual takes an LFT and receives a positive result. Given the test result, what is the conditional probability that they are infected?

(b) The individual then takes a PCR test, and receives a negative result. What is now the conditional probability that they are infected?

Mention any further assumptions you need to make in carrying out your calculations.

6. **(Euler's formula for the Riemann zeta function)**. For real numbers $s > 1$, the Riemann zeta function is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Fix $s > 1$ and consider a positive integer-valued random variable X with probability mass function given by

$$p_X(n) = \mathbb{P}(X = n) = \frac{1}{n^s}\frac{1}{\zeta(s)}, \quad n \geq 1.$$

(a) Let $k \geq 2$ be an integer. What is the probability that X is divisible by k?

(b) Let D_k be the event that X is divisible by k. Show that the events $\{D_p : p \text{ prime}\}$ are independent. *Hint: make sure you look back at the definition of independence for an infinite collection of events.*

(c) Recall from lectures that if the family $\{A_i, i \in I\}$ of events is independent, then so is the family $\{A_i^c, i \in I\}$. (Optional exercise: prove this result!) Use (b) to prove Euler's formula: for every $s > 1$,

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1},$$

where the product is over all prime numbers (one is not included as a prime). *Hint: you may assume that for independent events $A_1, A_2, \ldots$, we have $\mathbb{P}\left(\cap_{i=1}^{\infty} A_i\right) = \prod_{i=1}^{\infty} \mathbb{P}(A_i)$. Note that this does not follow directly from the definition of independence (proving it requires a slightly tricky argument using the countable additivity axiom).*