Problem set 1

1. (a) If $\mathcal{F}$ is a σ-algebra and $A_1, A_2, \ldots$ are in $\mathcal{F}$, then show that so is $\cap_{n=1}^{\infty} A_n$. (b) If A is a subset of B, written $A \subseteq B$, then show that $P(A) \leq P(B)$. (c) Show that $\emptyset \in \mathcal{F}$ and $P(\emptyset) = 0$, where $\emptyset$ denotes the empty set.

2. Let A_n, $n = 1, 2, \ldots$ be an increasing sequence of sets, i.e., $A_1 \subseteq A_2 \subseteq \cdots$. Let A be defined as the union of all the A_n. Then show that $P(A) = \lim_{n\to\infty} P(A_n)$. State and prove an analogous result for a decreasing sequence of sets.

3. Alice and Bob play a tennis match. Alice has probability p of winning a game if she is serving and q if each is receiving. (So the corresponding probabilities for Bob are $1 - q$ and $1 - p$). The first person to win 3 games wins the match.

There are two possible serving conventions; in one, serves alternate after each game and, in the other, the winner of a game gets to serve in the next game. If Alice starts serving, show without direct calculation that her probability of winning the match is the same under either convention.

(Hint: define a sample space by extending the match until Alice has served thrice and Bob has served twice, even if the outcome is already decided. Show that the event of Alice winning is the same under either convention, i.e., that it contains the same elements of the sample space.)

4. The gene that determines height in pea plants has two alleles, T and t. Plants with genotype TT and Tt are tall, while the tt genotype produces dwarf plants. (The allele T is called dominant as it determines the phenotype whenever it is present.) Suppose two heterozygous (i.e., having genotype Tt) tall plants are crossed, and the resultant offspring is tall. What is the probability that it is heterozygous?

5. (a) If $P(A)$ is either zero or one, then show that A is independent of B,

for any event B. (b) Show that, if $A_1, A_2 \ldots, A_n$ are mutually independent, then so are $A_1^c, A_2, \ldots, A_n$. Convince me, using a single word, that the same is true if any sub-collection of the sets A_i are complemented. (c) Show that $P(A \cap B|C) = P(A|C)P(B|A \cap C)$. Infer that

$$P\Big(\cap_{k=1}^n A_k\Big) = P(A_1) \prod_{k=2}^n P\Big(A_k \;\Big|\; \cap_{j=1}^{k-1} A_j\Big).$$

6. (from Ross(2006)). A blood test is 95% effective in detecting a certain disease but yields a false positive 1% of the time. If 0.5% of the population actually has the disease, what is the probability that a person has the disease given that the test result is positive? What if 5% of the population actually has the disease?

7. Sally Clark was convicted of the murder of her two children who both died suddenly of no apparent cause. The probability of a child dying of SIDS (sudden infant death syndrome) in the UK is about 1 in 8500. The prosecution argued that the probability of both children dying of SIDS is less than 1 in 73 million. Discuss this calculation, and also whether, by itself, it constitutes overwhelming evidence of guilt.

8. Birthday paradox: Suppose there are n days in a year and people are equally likely to be born on any given day. In a group of k people, what is the probability that at least two people share a birthday? Suppose n is very large. Roughly how large does k have to be (as a function of n) for this probability to be at least a half? (Hint: You might want to use the fact that $\log(1 + x) \approx x$ for small x.)

Here is another, heuristic, way to do the above calculation. What is the probability that a given pair of people share a birthday? Call this α. How many pairs of people are there? Call this N. Then, we "intuitively expect" that the number of pairs who share a birthday is approximately Poisson($N\alpha$). What is the probability that this Poisson random variable takes a value bigger than zero? Compare this with the answer obtained above.

The second calculation is a heuristic because we haven't shown that the number of pairs is indeed close to a Poisson random variable. We have also neglected triples, quadruples etc. of shared birthdays. But the heuristic can be made rigorous. Such heuristics play an important role in studying more complicated models than the birthday paradox, where exact calculations may not be feasible.

9. Compute the cumulative distribution functions of the Bernoulli and Geometric random variables.

10. For $\alpha > 0$, define $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$. If α is an integer, show that $\Gamma(\alpha) = (\alpha - 1)!$. Note that $0! = 1$. (Hint: Use integration by parts and induction.)