Exercise 1

(solutions to be handed in to Mariya Teteryatnikova one day before class)

1. Suppose the probability of an individual being born on any particular day of the year is given by 1/365.

(a) What is the probability that 2 people meeting at random have the same birthday?

(b) Suppose now that a group has 3 individuals. What is the probability that at least two of these individuals will share a birthday? What if the group has 4 individuals?

(c) How large must a group be such that the probability of finding at least two people with the same birthday is close to 50% (for this you will need to obtain an expression for the probability of at least two people sharing a birthday for an arbitrary group of size n.

2. A coin is to be tossed as many times as necessary to turn up one head. Thus the elements of c of the sample space are H, TH, TTH, $TTTH$ and so forth. Let the probability set function P assign to these elements the respective probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ etc. Show that $P(C) = 1$. Let $C_1 = \{c : c$ is $H, TH, TTH, TTTH$ or $TTTTH\}$. Compute $P(C_1)$. Next, suppose that $C_2 = \{c : c \text{ is } TTTTH \text{ or } TTTTTH\}$. Compute $P(C_2)$, $P(C_1 \cap C_2)$, $P(C_1 \cup C_2)$.

3. Examine whether functions $F_1$and F_2 are valid distribution functions. Explain your answers.

$$F_1(x) = \begin{cases} 0 & x < 0 \\ 2x^2 - \frac{2}{3}x^3 - x & 0 \leq x \leq 2 \\ 1 & 2 < x \end{cases}$$

$$F_2(x) = \begin{cases} 0 & x < 0 \\ 2x^2 - \frac{2}{3}x^3 & 0 \leq x \leq 2 \\ 1 & 2 < x \end{cases}$$

4. Suppose a certain accident annually kills 0.005% of the population. An insurance company provides insurance for 25,000 individuals. What is the probability that in a given year more than 5 of the insured individuals die due to the accident? Can you work out an estimate of the population mean death rate? How good an estimate is this likely to be? Can you provide a confidence interval for this estimate? [Justify briefly all your modelling assumptions in answering this question.]

5. Suppose X is a Poisson random variable with an (average) arrival rate of λ counts (occurences of a particular event, such as the arrival of a bus) per unit time. Thus,$E(X) = \lambda$. The probability of seeing a particular number k of counts in *one* unit of time is $P(X = k) = \lambda^k e^{-\lambda}/k!, k = 0, 1, 2, 3...$

For the same random variable, if we want to answer a question that involves an interval of $t > 0$ units of time rather than 1 unit of time, then we have a new random variable X_t that is Poisson distributed with mean λt. Now the probability of seeing exactly k counts in t units of time is $P(X_t = k) = (\lambda t)^k e^{-\lambda t}/k!$.

(a) Now consider a Poisson process with rate λ per unit time and the random variable W, which is the time one must wait to see the first count.

Explain briefly why the following two events are equivalent: $\{W > t\} = \{X_t = 0\}$

(b) Show then that the cumulative distribution of $W = P(W < t) = 1 - e^{-\lambda t}$ and also work out the density function of W. This is the so-called exponential density with rate λ. (Use the complement rule.)

(c) Can you generalise the argument to work out the density function of the random variable Z for the time one must wait to see the n^{th} count?

6. The table below shows the number of newborn boys and girls in the UK in 2003 and 2004.

(a) Set up a Bernoulli model for the 2003 data and estimate (by maximum likelihood) the parameter p giving the probability of a male birth.

(b) Consider a joint model for the data for 2003 and 2004, where p can vary across the two years (denote this by p_1 and p_2 respectively), and where all the observations are independent. Assuming that the joint likelihood is found by multiplying the two marginal likelihoods for 2003 and 2004, estimate p_1 and p_2.

(c) Finally, formulate (within the likelihood framework) a test of the null hypothesis $p_1 = p_2$ and test the hypothesis by computing a likelihood ratio statistic and comparing with a $\chi^2(1)$ density function.

7.

(a) Defining $B(n, k, p) = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}; k = 0, 1, 2, ...n$, show that $\lim_{n \to \infty;\ np=\lambda} B(n, k, p) = \frac{e^{-\lambda}\lambda^k}{k!}$.

(b) Derive the m.g.f. for a randon variable with probability distribution function $B(n, k, p)$ and show it is equal to $M_{B(n,k,p)}(t) = ((1-p) + pe^t)^n$.

(c) Work out the m.g.f. of a Poisson distribution with parameter λ, and show that under the same limiting process as in (a) above, the m.g.f. derived in(b) coincides with the m.g.f. for the Poisson.