# Probability for Statistics
## Problem Sheet 2

The first two questions are for you to practise working with the key definitions of probability functions, sigma algebras, and random variables, and should be accessible using material from week 2. Questions 3-5 are to give plenty of practice working with probability mass and density functions, cumulative distribution functions and transformations of random variables. Question 6 is an opportunity to think about the properties of a probability distribution. The final questions, for discussion, explore properties of waiting time distributions. For questions 3 onwards, it will be useful to have watched the week 3 videos, although some ideas will be familiar from last year.

1. Suppose $P$ and $Q$ are two probability functions defined on the same sample space $\Omega$ and sigma algebra $\mathcal{F}$.

   (a) Show that if $P(A) = Q(A)$ for all $A \in \mathcal{F}$ such that $P(A) \leq \frac{1}{2}$, then in fact $P = Q$ on all of $\mathcal{F}$.

   (b) Show by means of an explicit example that if instead we only have $P(A) = Q(A)$ for all $A \in \mathcal{F}$ such that $P(A) < \frac{1}{2}$, then $P$ and $Q$ need not agree on all of $\mathcal{F}$.

2. Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space and let $X$ and $Y$ be random variables with respect to $\mathcal{F}$. If $A \in \mathcal{F}$, define $Z : \Omega \to \mathbf{R}$ by

$$Z(\omega) = \begin{cases} X(\omega) & \omega \in A \\ Y(\omega) & \omega \notin A. \end{cases}$$

   (a) Show that $Z$ is a random variable with respect to $\mathcal{F}$.

   (b) Show that if instead $A \subseteq \Omega$ is not an event, i.e. $A \notin \mathcal{F}$, $Z$ need not be a random variable.

3. Suppose that $X$ is an absolutely continuous random variable with density function given by

$$f_X(x) = 4x^3, \text{ for } 0 < x < 1,$$

   and zero otherwise. Find the density functions of the following random variables:

   (a) $Y = X^4$, (b) $W = e^X$, (c) $Z = \log X$, (d) $U = (X - 0.5)^2$.

4. The measured radius of a circle, $R$, is an absolutely continuous random variable with density function given by

$$f_R(r) = 6r(1 - r), \text{ for } 0 < r < 1,$$

   and zero otherwise. Find the density functions of (a) the circumference and (b) the area of the circle.

5. Suppose that $X$ is an absolutely continuous random variable with density function given by

$$f_X(x) = \frac{\alpha}{\beta}\left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)}, \text{ for } x > 0,$$

   and zero elsewhere, with $\alpha$ and $\beta$ non-negative parameters.

(a) Find the density function and cdf of the random variable defined by $Y = \log X$.

(b) Find the density function of the random variable defined by $Z = \xi + \theta Y$.

6. On the probability space $(\Omega, \mathcal{F}, \Pr)$, let $Z$ be a random variable such that $\Pr(Z > 0) > 0$. Explain carefully why there exists $\delta > 0$ such that $\Pr(Z \geq \delta) > 0$.

**For discussion**

7. In this question, we look what happens to the geometric distribution when we pass from discrete to continuous time. Let $T$ have the waiting time geometric distribution with parameter $p$, so that

$$\Pr(T \geq j) = (1-p)^j, \qquad j = 0, 1, 2, \ldots$$

We think of $T$, which takes non-negative integer values, as the number of units of time we need to wait for an event to occcur. When $p$ is very small, $T$ typically takes very large values, so we seek to rescale time, so that the waiting times are given in more reasonable units. Let $M$ be a large number, such that $a = pM$ and $t = \frac{j}{M}$ are both small relative to $M$. What is the distribution of $U = \frac{T}{M}$, in terms of the parameter $a$? What important property has been preserved in this limit?

8. (*Harder*) Let $X_1$, $X_2$, $X_3$ be independent random variables, each with the mass function

$$\Pr(X_i = x) = (1-p_i)p_i^{x-1}, \qquad x = 1, 2, 3, \ldots$$

Show that

$$\Pr(X_1 < X_2 < X_3) = \frac{(1-p_1)(1-p_2)p_2 p_3^2}{(1-p_2p_3)(1-p_1p_2p_3)}.$$