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 from 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 Ω 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 $\mathcal{F}$.

   Objective: to practise working with probability functions in general, using only the properties specified by the axioms

   (a) Suppose $A \in \mathcal{F}$. If $P(A) \leq \frac{1}{2}$ then certainly $P(A) = Q(A)$ by hypothesis. If not, then $P(A^c) = 1 - P(A) < \frac{1}{2}$ so $P(A^c) = Q(A^c)$, so $1 - P(A) = 1 - Q(A)$, giving $P(A) = Q(A)$.

   (b) Take $\Omega = \{0, 1\}$ and $\mathcal{F} = \{\emptyset, \{0\}, \{1\}, \Omega\}$. Define

   $$P(\{0\}) = P(\{1\}) = \frac{1}{2}$$

   and

   $$Q(\{0\}) = \frac{1}{3} \qquad Q(\{1\}) = \frac{2}{3},$$

   with

   $$P(\emptyset) = Q(\emptyset) = 0, \qquad P(\Omega) = Q(\Omega) = 1.$$

   Reflect: perhaps surprisingly, there need not be any non-empty events with probability smaller than $\frac{1}{2}$.

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.

   Objective: to understand how to manufacture new random variables from old, working only with the definitions

(a) *Let $B \in \mathcal{B}$, then*

$$Z^{-1}(B) = \{\omega \in \Omega : Z(\omega) \in B\} = \{\omega \in A : X(\omega) \in B\} \cup \{\omega \in A^c : Y(\omega) \in B\}$$
$$= \left(X^{-1}(B) \cap A\right) \cup \left(Y^{-1}(B) \cap A^c\right).$$

Since $A, A^c \in \mathcal{F}$ and $X^{-1}(B), Y^{-1}(B) \in \mathcal{F}$, we see that $Z^{-1}(B) \in \mathcal{F}$.

(b) *Consider $\Omega = \{0, 1\}$ and the trivial sigma algebra $\mathcal{F} = \{\emptyset, \Omega\}$. Let $X, Y : \Omega \to \mathbf{R}$ be the constant random variables $X(\omega) = 0$ and $Y(\omega) = 1$. Certainly, these functions are random variables with respect to $\mathcal{F}$. But if $A = \{0\} \notin \mathcal{F}$, we see that $Z(0) = X(0) = 0$ and $Z(1) = Y(1) = 1$. But then Z is not a random variable with respect to $\mathcal{F}$, since $Z^{-1}(\{0\}) = \{0\} \notin \mathcal{F}$.*

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

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

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) $U = (X - 0.5)^2$.

The cdf of X, F_X is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = \int_0^x 4t^3\,\mathrm{d}t = x^4, \qquad 0 < x < 1.$$

(a) *$Y = X^4$, so $\mathbb{Y} = (0, 1)$, and from first principles, for $y \in \mathbb{Y}$,*

$$F_Y(y) = \Pr(Y \le y) = \Pr(X^4 \le y) = \Pr(X \le y^{1/4}) = F_X(y^{1/4}) = y.$$

Thus, $f_Y(y) = 1$, for $0 < y < 1$.

(b) *$W = e^X$, so $\mathbb{W} = (1, e)$, and from first principles, for $w \in \mathbb{W}$,*

$$F_W(w) = \Pr(W \le w) = \Pr(e^X \le w) = \Pr(X \le \log w) = F_X(\log w) = (\log w)^4$$

$$\implies f_W(w) = \frac{4(\log w)^3}{w}, \qquad 1 < w < e.$$

(c) *$Z = \log X$, so $\mathbb{Z} = (-\infty, 0)$, and from first principles, for $z \in \mathbb{Z}$,*

$$F_Z(z) = \Pr(Z \le z) = \Pr(\log X \le z) = \Pr(X \le e^z) = F_X(e^z) = e^{4z}.$$

Thus, $f_Z(z) = 4e^{4z}$, for $-\infty < z < 0$.

(d) *$U = (X - 0.5)^2$, so $\mathbb{U} = (0, 0.25)$, and from first principles, for $u \in \mathbb{U}$,*

$$F_U(u) = \Pr(U \le u) = \Pr[\,(X - 0.5)^2 \le u\,] = \Pr(-\sqrt{u} + 0.5 \le X \le \sqrt{u} + 0.5)$$

$$= F_X(\sqrt{u} + 0.5) - F_X(-\sqrt{u} + 0.5) = (0.5 + \sqrt{u})^4 - (0.5 - \sqrt{u})^4$$

$$\implies f_U(u) = \frac{2}{\sqrt{u}}\left[(0.5 + \sqrt{u})^3 + (0.5 - \sqrt{u})^3\right] = \frac{1 + 12u}{2\sqrt{u}}, \qquad 0 < u < 0.25.$$

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.

We have $f_R(r) = 6r(1-r)$*, for* $0 < r < 1$*, and hence*

$$F_R(r) = r^2(3-2r), \quad 0 < r < 1.$$

(a) *Circumference:* $Y = 2\pi R$*, so* $\mathbb{Y} = (0, 2\pi)$*, and from first principles, for* $y \in \mathbb{Y}$*,*

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) = \Pr(2\pi R \le y) = \Pr(R \le y/2\pi) = F_R(y/2\pi) = \frac{3y^2}{4\pi^2} - \frac{2y^3}{8\pi^3} \\ \Longrightarrow f_Y(y) &= \frac{6y}{8\pi^3}(2\pi - y), \quad 0 < y < 2\pi. \end{aligned}$$

(b) *Area:* $Z = \pi R^2$*, so* $\mathbb{Z} = (0, \pi)$*, and from first principles, for* $z \in \mathbb{Z}$*, recalling that* f_R *is only positive when* $0 < z < \pi$*,*

$$\begin{aligned} F_Z(z) &= \Pr(Z \le z) = \Pr(\pi R^2 \le z) = \Pr(R \le \sqrt{z/\pi}) = F_R(\sqrt{z/\pi}) = \frac{3z}{\pi} - 2\left\{\frac{z}{\pi}\right\}^{3/2} \\ \Longrightarrow f_Z(z) &= 3\pi^{-3/2}(\sqrt{\pi} - \sqrt{z}), \quad 0 < z < \pi. \end{aligned}$$

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 α and β 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 $Z = \xi + \theta Y$.

By integration

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = \int_0^x \frac{\alpha}{\beta}\left(\frac{\beta}{\beta+t}\right)^{\alpha+1} \mathrm{d}t = -\left(\frac{\beta}{\beta+t}\right)^{\alpha}\Bigg|_0^x = 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha}, \quad x > 0.$$

(a) *If* $Y = \log X$*, then* $\mathbb{Y} = \mathbb{R}$*, and*

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) = \Pr(\log X \le y) = \Pr(X \le e^y) = F_X(e^y) = 1 - \left(1 + \frac{e^y}{\beta}\right)^{-\alpha} \\ \Longrightarrow f_Y(y) &= \frac{\alpha}{\beta}e^y\left(\frac{\beta}{\beta+e^y}\right)^{\alpha+1}, \quad y \in \mathbb{R}. \end{aligned}$$

(b) *If* $Z = \xi + \theta Y$*, then* $Y = (Z - \xi)/\theta$*, so the density of* Z *can be found easily using transformation techniques*

$$f_Z(z) = \frac{\alpha}{\beta}e^{(z-\xi)/\theta}\left(\frac{\beta}{\beta + e^{(z-\xi)/\theta}}\right)^{\alpha+1}\frac{1}{|\theta|}, \text{ for } z \in \mathbb{R}.$$

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 $n \geq 1$, define $A_n = \{Z \geq \frac{1}{n}\}$. Then $A_1 \subseteq A_2 \subseteq \ldots$ is an increasing sequence of events, and

$$\{Z > 0\} = \bigcup_{n=1}^{\infty} A_n.$$

   Applying the continuity property of $\Pr$, this then gives ,

$$\Pr(\{Z > 0\}) = \Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \Pr(A_n).$$

   If it were the case that $\Pr(A_n) = 0$ for all $n \geq 1$, then the right hand limit would be 0, which is a contradiction. Hence there exists $n \geq 1$ such that $\Pr(A_n) > 0$, so taking $\delta = \frac{1}{n}$ suffices.

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?

$$\Pr(U \geq u) = \Pr\left(\frac{T}{M} \geq u\right) = \Pr(T \geq uM) = (1-p)^{uM} = \left(1 - \frac{a}{M}\right)^{uM} \approx e^{-au}.$$

   So U has the exponential distribution with rate a. This limiting distribution inherits the memoryless property.

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

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

   One can do this directly by evaluating the sum

$$
\begin{aligned}
Pr(X_1 < X_2 < X_3) &= \sum_{1 \le i < j < k < \infty} (1-p_1)(1-p_2)(1-p_3) p_1^{i-1} p_2^{j-1} p_3^{k-1} \\
&= (1-p_1)(1-p_2) \sum_{1 \le i < j < \infty} p_1^{i-1} p_2^{j-1} p_3^{j}. \\
&= (1-p_1)(1-p_2) p_3 \sum_{1 \le i < \infty} \frac{p_1^{i-1} (p_2 p_3)^i}{1 - p_2 p_3} \\
&= \frac{(1-p_1)(1-p_2) p_2 p_3^2}{(1-p_2 p_3)(1-p_1 p_2 p_3)}.
\end{aligned}
$$

An alternative approach uses the properties of the minimum of two geometric random variables.

First we compute $\Pr(X < Y)$ *where* X *and* Y *are independent geometric random variables with success probabilities* $1-p$ *and* $1-q$*, respectively.*

$$
\begin{aligned}
\Pr(X < Y) &= \sum_{k=1}^{\infty} \Pr(X < Y | Y = k) \Pr(Y = k) = \sum_{k=2}^{\infty} \left(1 - p^{k-1}\right)(1-q) q^{k-1} \\
&= (1-q) \sum_{k=2}^{\infty} q^{k-1} - (1-q) \sum_{k=2}^{\infty} (pq)^{k-1} = q - \frac{(1-q)pq}{1-pq} = \frac{q(1-p)}{1-pq}.
\end{aligned}
$$

Now, note that $\min(X, Y)$ *is a geometric variable with failure probability* pq*, since*

$$
\Pr(\min(X, Y) > k) = \Pr(X > k \cap Y > k) = \Pr(X > k) \Pr(Y > k) = p^k q^k.
$$

Giving

$$
\Pr(\min(X, Y) \le k) = 1 - (pq)^k,
$$

which is the CDF of a geometric random variable.

So then now

$$
\Pr(X_1 < X_2 < X_3) = \Pr(X_1 < X_2 \cap X_2 < X_3) = \Pr(X_1 < X_2 | X_2 < X_3) \Pr(X_2 < X_3).
$$

From the above calculation, we see that

$$
\Pr(X_2 < X_3) = \frac{p_3(1-p_2)}{1-p_2 p_3}
$$

Now, conditional on $X_2 < X_3$*,* X_2 *is the minimum of two geometric random variables, so by the result above,*

$$
\Pr(X_1 < X_2 | X_2 < X_3) = \frac{p_2 p_3 (1-p_1)}{1 - p_1 p_2 p_3},
$$

giving the same result.