

# Probability for Statistics

## Problem Sheet 3

Questions marked (†) may require material from Thursdays lecture.

1. 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:

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

The cdf of  $X$ ,  $F_X$  is given by

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

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

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

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

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

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

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

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

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

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  $u \in \mathbb{U}$ ,

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

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

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

2. 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  $y \in \mathbb{Y}$ ,

$$F_Y(y) = \Pr(Y \leq y) = \Pr(2\pi R \leq y) = \Pr(R \leq y/2\pi) = F_R(y/2\pi) = \frac{3y^2}{4\pi^2} - \frac{2y^3}{8\pi^3}$$

$$\implies f_Y(y) = \frac{6y}{8\pi^3}(2\pi - y), \quad 0 < y < 2\pi.$$

(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  $0 < z < \pi$ ,

$$F_Z(z) = \Pr(Z \leq z) = \Pr(\pi R^2 \leq z) = \Pr(R \leq \sqrt{z/\pi}) = F_R(\sqrt{z/\pi}) = \frac{3z}{\pi} - 2 \left\{ \frac{z}{\pi} \right\}^{3/2}$$

$$\implies f_Z(z) = 3\pi^{-3/2}(\sqrt{\pi} - \sqrt{z}), \quad 0 < z < \pi.$$

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

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

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  $Z = \xi + \theta Y$ .

By integration

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

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

$$F_Y(y) = \Pr(Y \leq y) = \Pr(\log X \leq y) = \Pr(X \leq e^y) = F_X(e^y) = 1 - \left(1 + \frac{e^y}{\beta}\right)^{-\alpha}$$

$$\implies f_Y(y) = \frac{\alpha}{\beta} e^y \left(\frac{\beta}{\beta + e^y}\right)^{\alpha+1}, \quad y \in \mathbb{R}.$$

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

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

4. Let  $X$  be an absolutely continuous random variable with range  $\mathbb{X} = \mathbb{R}^+$ , pdf  $f_X$  and cdf  $F_X$ .

(a) Show that

$$\mathbb{E}(X) = \int_0^{\infty} [1 - F_X(x)] dx.$$

(b) Show also that for integer  $r \geq 1$ ,

$$\mathbb{E}(X^r) = \int_0^{\infty} r x^{r-1} [1 - F_X(x)] dx.$$

(c) Find a similar expression for  $E(X^r)$  for random variables for which  $\mathbb{X} = \mathbb{R}$ .

(a)

$$\begin{aligned} E(X) &= \int_0^\infty x f_X(x) dx = \int_0^\infty \left\{ \int_0^x dy \right\} f_X(x) dx = \int_0^\infty \left\{ \int_y^\infty f_X(x) dx \right\} dy \quad (1) \\ &= \int_0^\infty (1 - F_X(y)) dy \equiv \int_0^\infty (1 - F_X(x)) dx. \end{aligned}$$

**Reflect:** Notice that the change in the range of integration in the third equality follows from the change in the order of integration. The exchange of order of integration is valid if we know that the expectation integral is finite. The result also holds in the discrete case with integrals replaced by summations. The important thing is to remember the trick of introducing a second integral involving dummy variable  $y$ . The rest of the result follows after careful manipulation of the double integral. Intuitively, to determine the limits when changing the order of integration, note that in (1), we are integrating over  $(x, y)$  such that  $0 \leq x < \infty, 0 \leq y < x$ , or equivalently,  $0 \leq y \leq x < \infty$ . From this, we see the constraints on  $y$  for a fixed  $x$  are  $0 \leq y < x$ , and the constraints on  $x$  for a fixed  $y$  are  $y \leq x < \infty$ . (We can also determine the ranges by sketching the region  $0 \leq y \leq x < \infty$ ).

Many people get quite excited about this result, and there are some nice articles giving geometric interpretations etc. See e.g. [Demystifying the Integrated Tail Probability Expectation Formula by Ambrose Lo](#) (Shibboleth login probably needed).

(b)

$$\begin{aligned} E(X^r) &= \int_0^\infty x^r f_X(x) dx = \int_0^\infty \left\{ \int_0^x r y^{r-1} dy \right\} f_X(x) dx = \int_0^\infty \left\{ \int_y^\infty f_X(x) dx \right\} r y^{r-1} dy \\ &= \int_0^\infty (1 - F_X(y)) r y^{r-1} dy \equiv \int_0^\infty r x^{r-1} (1 - F_X(x)) dx. \end{aligned}$$

(c) For a random variable that takes values on  $\mathbb{R}$ , we split the integral into two at the origin and proceed as above, as follows:

$$\begin{aligned} E(X^r) &= \int_{-\infty}^\infty x^r f_X(x) dx = \int_{-\infty}^0 x^r f_X(x) dx + \int_0^\infty x^r f_X(x) dx \\ &= \int_{-\infty}^0 \left\{ \int_0^x r y^{r-1} dy \right\} f_X(x) dx + \int_0^\infty r x^{r-1} (1 - F_X(x)) dx \\ &= \int_{-\infty}^0 \left\{ - \int_x^0 r y^{r-1} dy \right\} f_X(x) dx + \int_0^\infty (1 - F_X(y)) r y^{r-1} dy \\ &= - \int_{-\infty}^0 r y^{r-1} \left\{ \int_{-\infty}^y f_X(x) dx \right\} dy + \int_0^\infty (1 - F_X(y)) r y^{r-1} dy \\ &= - \int_{-\infty}^0 r y^{r-1} F_X(y) dy + \int_0^\infty (1 - F_X(y)) r y^{r-1} dy. \end{aligned}$$

5. Consider two absolutely continuous random variables  $X$  and  $Y$  such that

$$\Pr(X \leq x \text{ and } Y \leq y) = (1 - e^{-x}) \left( \frac{1}{2} + \frac{1}{\pi} \tan^{-1} y \right), \text{ for } x > 0 \text{ and } -\infty < y < \infty,$$

with

$$\Pr(X \leq x \text{ and } Y \leq y) = 0, \text{ for } x \leq 0.$$

Find the joint pdf,  $f_{X,Y}$ . Are  $X$  and  $Y$  independent? Justify your answer.

Let  $F_{X,Y}(x, y) = \Pr(X \leq x \text{ and } Y \leq y)$ . By the fundamental theorem of calculus, the function

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial t_1 \partial t_2} F_{X,Y}(t_1, t_2) \Big|_{t_1=x, t_2=y} = \frac{e^{-x}}{\pi(1+y^2)}$$

is a pdf of  $(X, Y)$ . That is to say, probabilities of measurable regions,  $\Pr((X, Y) \in \mathcal{R})$ , can be computed by integrating  $f_{X,Y}$  over  $\mathcal{R}$ . Finally, because (i)  $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ , and (ii) the support of  $(X, Y)$  is  $\mathbf{R}^+ \times \mathbf{R}$ ,  $X$  and  $Y$  are independent.

6. (†) Suppose that the joint pdf of  $X$  and  $Y$  is given by

$$f_{X,Y}(x, y) = 24xy, \text{ for } x > 0, y > 0, \text{ and } x + y < 1,$$

and zero otherwise. Find

- (a) the marginal pdf of  $X$ ,  $f_X$ ,
- (b) the marginal pdf of  $Y$ ,  $f_Y$ ,
- (c) the conditional pdf of  $X$  given  $Y = y$ ,  $f_{X|Y}$ ,
- (d) the conditional pdf of  $Y$  given  $X = x$ ,  $f_{Y|X}$ ,
- (e) the expected value of  $X$ ,
- (f) the expected value of  $Y$ ,
- (g) the conditional expected value of  $X$  given  $Y = y$ , and
- (h) the conditional expected value of  $Y$  given  $X = x$ .

[Hint: Sketch the region on which the joint density is non-zero; remember that the integrand is only non-zero for some part of the integral range.]

- (a) The joint pdf of  $X$  and  $Y$  is given by

$$f_{X,Y}(x, y) = 24xy, \text{ for } x > 0, y > 0, x + y < 1,$$

and zero otherwise, the marginal pdf  $f_X$  is given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \int_0^{1-x} 24xy dy = 24x \left[ \frac{y^2}{2} \right]_0^{1-x} = 12x(1-x)^2, \text{ for } 0 < x < 1,$$

as the integrand is only non-zero when  $0 < x + y < 1 \implies 0 < y < 1 - x$  for fixed  $x$ .

- (b) Because  $f_{X,Y}(x, y)$  and its support are symmetric in  $x$  and  $y$ , the marginal densities are the same,

$$f_Y(y) = 12y(1-y)^2, \text{ for } 0 < y < 1.$$

(c) The conditional density is proportional to the joint density,

$$f_{X|Y}(x|y) \propto f_{X,Y}(x,y) = 24xy = cxy \text{ for } 0 < x < 1-y.$$

To find  $c$ ,

$$\int_0^{1-y} xy \, dx = \frac{x^2 y}{2} \Big|_0^{1-y} = \frac{(1-y)^2 y}{2}.$$

So  $c = 2/y(1-y)^2$  and

$$f_{X|Y}(x|y) = \frac{2x}{(1-y)^2} \text{ for } 0 < x < 1-y.$$

(d) Because  $f_{X,Y}(x,y)$  and its support are symmetric in  $x$  and  $y$ , the conditional densities are also symmetric,

$$f_{Y|X}(y|x) = \frac{2y}{(1-x)^2} \text{ for } 0 < y < 1-x.$$

(e) & (f) Because  $X$  and  $Y$  have the same marginal distribution, they have the same expectation,

$$\begin{aligned} E(Y) &= E(X) = \int_0^1 x [12x - 24x^2 + 12x^3] dx = \int_0^1 [12x^2 - 24x^3 + 12x^4] dx \\ &= \left[ 4x^3 - 6x^4 + \frac{12}{5}x^5 \right] \Big|_0^1 = 0.4. \end{aligned}$$

(g) & (h) Because  $X | Y = y$  and  $Y | X = x$  have the same conditional distribution, they have the same conditional expectation,

$$E(X | Y) = \int_0^{1-y} x f_{X|Y}(x | y) dx = \int_0^{1-y} \frac{2x^2}{(1-y)^2} dx = \frac{2x^3}{3(1-y)^2} \Big|_0^{1-y} = \frac{2}{3}(1-y)$$

$$\text{and } E(Y | X) = \frac{2}{3}(1-x).$$

### For discussion

7. (†) Consider two independent random variables  $X_1$  and  $X_2$ , exponentially distributed with rate 1. Suppose we wish to consider the density function of  $X_1$  conditional on the event  $\{X_1 = X_2\}$ .

(a) One way to do this is to consider the variable  $Z = X_1 - X_2$ , and condition on the event  $Z = 0$ . Find the pdf  $f(x_1 | z = 0)$ . Since  $X_1, X_2$  are independent exponential random variables with rate 1,  $f_{X_1, X_2}(x_1, x_2) = e^{-x_1 - x_2}$ . To find the joint distribution of  $X_1, Z$ , we can apply Proposition 3.7 (Change of variables for probability densities) with  $W = X_1, Z = X_1 - X_2$ .

In this case  $J = \left| \det \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \right| = 1$ . Hence,

$$f_{X_1, Z}(x_1, z) = e^{-(2x_1 - z)}, \quad x_1 > 0, z < x_1.$$

So then the value of the marginal pdf of  $Z$  at  $Z = 0$  is given by

$$f_Z(z) = \int_0^\infty f_{X_1, Z}(x_1, 0) dx_1 = \int_0^\infty e^{-2x_1} dx_1 = \frac{1}{2}.$$

Then we see

$$f(x_1 | z = 0) = \frac{f_{X_1, Z}(x_1, 0)}{f_Z(0)} = 2e^{-2x_1} \quad x_1 > 0.$$

- (b) Alternatively, one could consider the variable  $W = \frac{X_2}{X_1}$ , and condition on the event  $W = 1$ . Find the pdf  $f(x_1|w = 1)$ .

By standard transformation methods,

$$f_{X_1, W}(x_1, w) = x_1 e^{-(x_1 + wx_1)}, \quad x_1 > 0, w > 0.$$

So then the value of the marginal pdf of  $W$  at  $W = 1$  is given by

$$f_W(w) = \int_0^\infty f_{X_1, W}(x_1, 1) dx_1 = \int_0^\infty x_1 e^{-2x_1} dx_1 = \frac{1}{4}.$$

Then we see

$$f(x_1|w = 1) = \frac{f_{X_1, W}(x_1, 1)}{f_W(1)} = 4x_1 e^{-2x_1} \quad x_1 > 0.$$

- (c) Comment on your answers to the two parts above. (This is an instance of the Borel-Kolmogorov paradox.)

**Reflect: Conditioning on an event of probability zero (such as  $\{X_1 = X_2\}$ ) is not well-defined. Here (say for  $Z$ ), we condition on events  $A_n$  with  $\Pr(A_n) > 0$  but  $\lim_{n \rightarrow \infty} \Pr(A_n) = 0$ . When for  $W$  we approach the limit using a different sequence, unsurprisingly, we get a different answer.**

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

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

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

$$\begin{aligned} \Pr(X_1 < X_2 < X_3) &= \sum_{1 \leq i < j < k < \infty} (1 - p_1)(1 - p_2)(1 - p_3)p_1^{i-1}p_2^{j-1}p_3^{k-1} \\ &= (1 - p_1)(1 - p_2) \sum_{1 \leq i < j < \infty} p_1^{i-1}p_2^{j-1}p_3^j \\ &= (1 - p_1)(1 - p_2)p_3 \sum_{1 \leq i < \infty} \frac{p_1^{i-1}(p_2p_3)^i}{1 - p_2p_3} \\ &= \frac{(1 - p_1)(1 - p_2)p_2p_3^2}{(1 - p_2p_3)(1 - p_1p_2p_3)}. \end{aligned}$$

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.

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

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

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

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.