Probability for Statistics
Problem Sheet 3

The objective here is to practise working with joint distributions of random variables, and the quantities that can be derived from them. All questions should be accessible by the time you have watched all week 4 videos.

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

   (a) Show that
   $$\mathrm{E}(X) = \int_0^\infty [1 - F_X(x)] \, \mathrm{d}x.$$

   (b) Show also that for integer $r \geq 1$,
   $$\mathrm{E}(X^r) = \int_0^\infty r x^{r-1} [1 - F_X(x)] \, \mathrm{d}x.$$

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

   Objective: to derive an alternative (perhaps surprising) expression for expectation. Along the way, we'll make a connection with multiple integration.

   (a)

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

   Reflect: Notice that the change in the range of integration in the third equality follows from the change in the order or 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.

   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)

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

(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:*

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

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

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

with

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

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

Objective: practise working with joint distributions, and develop intuition for when random variables are independent .

Let $F_{X,Y}(x, y) = \Pr(X \le x$ and]$Y \le y)$. 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_{XY} 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.

3. Suppose that the joint pdf of X and Y is given by

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

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.]

Objective: gain familiarity with the quantities (e.g. marginal/conditional distributions) that can be computed from a joint distribution.

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

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

and zero otherwise, the marginal pdf f_X is given by

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

as the integrand is only non-zero when $0 < x + y < 1 \Longrightarrow 0 < y < 1 - x$ 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,

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

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

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

To find c,

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

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

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

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

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

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

$$\begin{aligned} \mathrm{E}(Y) \ &= \mathrm{E}(X) = \textstyle\int_0^1 x\big[12x - 24x^2 + 12x^3\big]\mathrm{d}x = \int_0^1 \big[12x^2 - 24x^3 + 12x^4\big]\mathrm{d}x \\ &= \left[4x^3 - 6x^4 + \tfrac{12}{5}x^5\right]\Bigg|_0^1 = 0.4. \end{aligned}$$

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

$$\mathrm{E}(X \mid Y) = \int_0^{1-y} x f_{X|Y}(x \mid y)\mathrm{d}x = \int_0^{1-y} \frac{2x^2}{(1-y)^2}\mathrm{d}x = \frac{2x^3}{3(1-y)^2}\Bigg|_0^{1-y} = \frac{2}{3}(1-y)$$

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

4. (Harder) Suppose that X and Y have joint pdf that is constant on the range $\mathbb{X}^{(2)} \equiv (0,1) \times (0,1)$, and zero otherwise. Find the marginal pdf of the random variables $U = X/Y$ and $V = -\log(XY)$, stating clearly the range of the transformed random variable in each case.

[Hint: For U, you might consider first the joint pdf of (U, X), then obtain the marginal pdf of U. For V, consider the joint pdf of $(V, -\log X)$, then obtain the marginal pdf of V. These choices result in much simpler calculations than those required to derive the joint transformation from (X, Y) to (U, V).]

Objective: Gain practise computing marginal distributions in a somewhat trickier case.

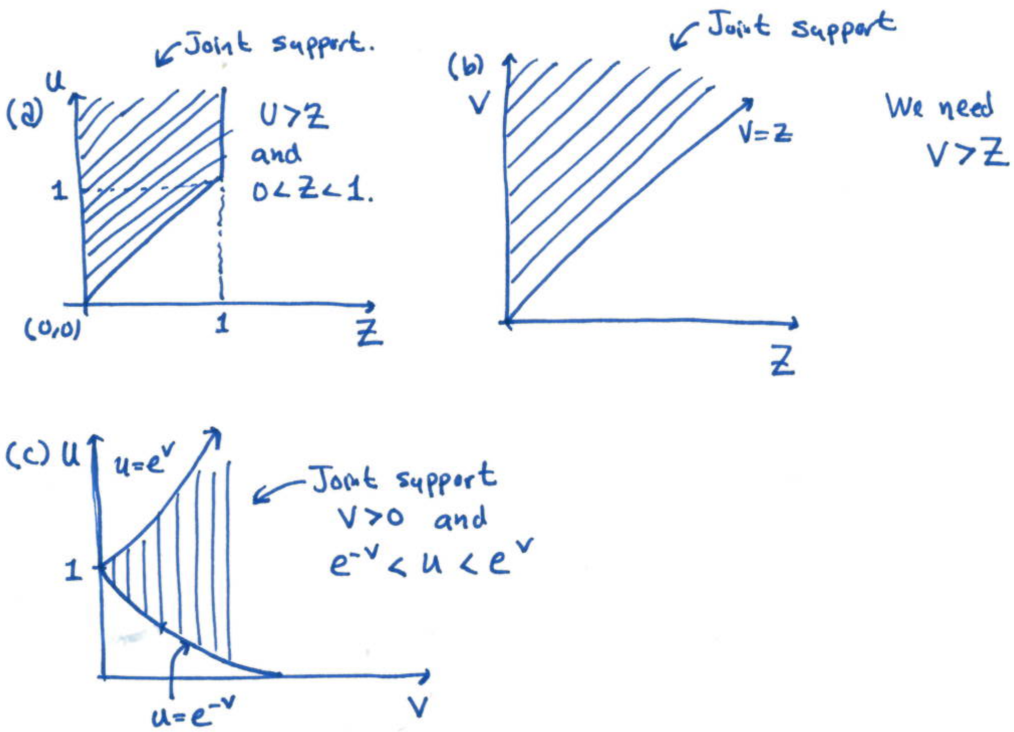

Figure 1: Sketches of the joint support of the transformed random variables in Problem 6.

First, put $U = X/Y$ and $Z = X$; the inverse transformations are therefore $X = Z$ and $Y = Z/U$. The first step is to sketch the support of the two new variables, see Figure 1(a) which shows how the new variables are constrained by $0 < Z < \min\{U, 1\}$ (because $Y < 1$). In terms of the multivariate transformation theorem, we have transformation functions defined by

$$g_1(t_1, t_2) = t_1/t_2, \qquad g_1^{-1}(t_1, t_2) = t_2,$$

$$g_2(t_1, t_2) = t_1, \qquad g_2^{-1}(t_1, t_2) = t_2/t_1,$$

and the Jacobian of the transformation is given by

$$J(u, z) = \begin{vmatrix} 0 & 1 \\ -z/u^2 & 1/u \end{vmatrix} = \frac{z}{u^2}$$

and hence

$$f_{U,Z}(u, z) = f_{X,Y}(z, z/u)\, z/u^2 = z/u^2, \qquad (u, z) \in \mathbb{U}^{(2)} \equiv \{(u, z) : 0 < z < \min\{u, 1\}, u > 0\},$$

and zero otherwise, and so

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,Z}(u,z)\,\mathrm{d}z = \int_0^{\min\{u,1\}} z/u^2\,\mathrm{d}z = \frac{(\min\{u,1\})^2}{2u^2}, \qquad u > 0.$$

Now, for V, put $V = -\log(XY)$ and $Z = -\log X$; the inverse transformations are therefore $X = e^{-Z}$ and $Y = e^{-(V-Z)}$. Note =that $0 < Z < V$, see Figure 1(b). In terms of the theorem, we have transformation functions defined by

$$g_1(t_1,t_2) = -\log(t_1 t_2), \qquad g_1^{-1}(t_1,t_2) = e^{-t_2},$$

$$g_2(t_1,t_2) = -\log t_1, \qquad g_2^{-1}(t_1,t_2) = e^{-(t_1-t_2)},$$

and the Jacobian of the transformation is given by

$$J(v,z) = \begin{vmatrix} 0 & -e^{-z} \\ -e^{-(v-z)} & e^{-(v-z)} \end{vmatrix} = e^{-v},$$

and hence

$$f_{V,Z}(v,z) = f_{X,Y}(e^{-z}, e^{-(v-z)})\, e^{-v} = e^{-v}, \qquad (v,z) \in \mathbb{V}^{(2)} \equiv \{(v,z) : 0 < z < v < \infty\},$$

and zero otherwise, and so

$$f_V(v) = \int_{-\infty}^{\infty} f_{V,Z}(v,z)\,\mathrm{d}z = \int_0^v e^{-v}\,\mathrm{d}z = ve^{-v}, \qquad v > 0,$$

and zero otherwise.

Reflect: of course, we should be able to obtain the same results by the joint transformation. Let's check. Set

$$\begin{array}{l} U = X/Y \\ V = -\log(XY) \end{array} \iff \begin{array}{l} X = U^{1/2}e^{-V/2} \\ Y = U^{-1/2}e^{-V/2} \end{array}$$

and note that, as X and Y lie in $(0,1)$ we have $XY < X/Y$ and $XY < Y/X$, giving constraints $e^{-V} < U$ and $e^{-V} < 1/U$, so that $0 < e^{-V} < \min\{U, 1/U\}$. The support of U and V is sketched in Figure 1(c). The Jacobian of the transformation is

$$J(u,v) = \begin{vmatrix} \dfrac{u^{-1/2}e^{-v/2}}{2} & -\dfrac{u^{1/2}e^{-v/2}}{2} \\ -\dfrac{u^{-3/2}e^{-v/2}}{2} & -\dfrac{u^{-1/2}e^{-v/2}}{2} \end{vmatrix} = u^{-1}e^{-v}/2.$$

Hence

$$f_{U,V}(u,v) = u^{-1}e^{-v}/2, \qquad 0 < e^{-v} < \min\{u, 1/u\},\ u > 0.$$

The corresponding marginals are given below: let $g(u) = -\log(\min\{u, 1/u\})$, then

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)\,\mathrm{d}v = \int_{g(u)}^{\infty} \frac{e^{-v}}{2u}\,\mathrm{d}v = -\frac{e^{-v}}{2u}\Bigg|_{g(u)}^{\infty} = \frac{\min\{u,1/u\}}{2u}, \quad u > 0,$$

$$f_V(v) = \int_{-\infty}^{\infty} f_{U,V}(u,v)\,\mathrm{d}u = \int_{e^{-v}}^{e^v} \frac{e^{-v}}{2u}\,\mathrm{d}u = \frac{\log u}{2}e^{-v}\Bigg|_{e^{-v}}^{e^v} = ve^{-v}, \qquad v > 0.$$

5. Suppose that X and Y are absolutely continuous random variables with pdf given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\left(x^2+y^2\right)\right\}, \text{ for } x,y \in \mathbb{R}.$$

(a) Let the random variable U be defined by $U = X/Y$. Find the pdf of U.

(b) Suppose now that $S \sim \chi^2_\nu$ is independent of X and Y. (The pdf of S is given by

$$f_S(s) = c(\nu)s^{\nu/2-1}e^{-s/2}, \text{ for } s > 0,$$

where ν is a positive integer and $c(\nu)$ is a normalizing constant depending on ν.) Find the pdf of random variable T defined by

$$T = \frac{X}{\sqrt{S/\nu}}.$$

This is the pdf of a t random variable with ν degrees of freedom.

Objective: Derive distributions of huge practical signifiance in statistical modelling. (a) is the Cauchy distribution, the most famous counter-example to the central limit theorem and (b) is the t distribution, the distribution of the sample mean of a normal random sample when standardized by an appropriate estimate of its standard error. In practical contents, the random variable labelled here as X is the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ of a random sample $(X_1,\ldots,X_n)$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$. Given that $\bar{X}$ and S^2 are both functions of the same random sample, it may be surprising that they are independent random variables. This fact is certainly not obvious - it follows from Cochran's theorem.

(a) Put $U = X/Y$ and $V = Y$; the inverse transformations are therefore $X = UV$ and $Y = V$. In terms of the multivariate transformation theorem, we have transformation functions defined by

$$g_1(t_1,t_2) = t_1/t_2, \qquad g_1^{-1}(t_1,t_2) = t_1t_2,$$

$$g_2(t_1,t_2) = t_2, \qquad g_2^{-1}(t_1,t_2) = t_2,$$

and the Jacobian of the transformation is given by

$$J(u,v) = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$$

and hence

$$f_{U,V}(u,v) = f_{X,Y}(uv,v)\ |v| = \left(\frac{1}{2\pi}\right)\exp\left\{-\frac{1}{2}(u^2v^2+v^2)\right\}|v|, \text{ for } (u,v) \in \mathbb{R}^2.$$

Now, for any real u,

$$\begin{aligned} f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)\,\mathrm{d}v &= \int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\right)\exp\left\{-\frac{1}{2}(u^2v^2+v^2)\right\}|v|\ \mathrm{d}v \\ &= \left(\frac{1}{\pi}\right)\int_0^{\infty} v\exp\left\{-\frac{v^2}{2}(1+u^2)\right\}\mathrm{d}v \quad \textit{(as integrand is even function)} \end{aligned}$$

$$= \left(\frac{1}{\pi}\right)\left[-\frac{1}{(1+u^2)}\exp\left\{-\frac{v^2}{2}(1+u^2)\right\}\right]\Bigg|_0^\infty = \frac{1}{\pi(1+u^2)},$$

with the final step following by direct integration.

Reflect: This is the infamous Cauchy distribution. It is most commonly encountered as a case where results such as the central limit theorem do not apply, because U does not have finite moments of any order. In particular, it does not have a finite mean or variance.

As its pdf is an even function, we might be tempted to assert that its mean should be zero. But this does not respect the definition we made when defining expectation. $E|U|$ is not finite, so $E(U)$ is not defined. Equivalently, its mean is zero only if we allow $\infty - \infty = 0$, which we do not.

(b) Now put $T = X/\sqrt{S/\nu}$ and $R = S$; the inverse transformations are therefore $X = T\sqrt{R/\nu}$ and $S = R$. In terms of the multivariate transformation theorem, we have transformation functions from $(X,S) \to (T,R)$ defined by

$$g_1(t_1,t_2) = t_1/\sqrt{t_2/\nu}, \qquad g_1^{-1}(t_1,t_2) = t_1\sqrt{t_2/\nu},$$

$$g_2(t_1,t_2) = t_2, \qquad g_2^{-1}(t_1,t_2) = t_2,$$

and the Jacobian of the transformation is given by

$$J(t,r) = \begin{vmatrix} \sqrt{\frac{r}{\nu}} & \frac{t}{2\sqrt{r\nu}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{r}{\nu}},$$

and hence

$$f_{T,R}(t,r) = f_{X,S}\left(t\sqrt{\frac{r}{\nu}}, r\right)\sqrt{\frac{r}{\nu}} = f_X\left(t\sqrt{\frac{r}{\nu}}\right) f_S(r)\sqrt{\frac{r}{\nu}}, \text{ for } t \in \mathbb{R}, r \in \mathbb{R}^+,$$

and zero otherwise. Now, for any real t,

$$\begin{aligned}
f_T(t) &= \int_{-\infty}^{\infty} f_{T,R}(t,r)\,\mathrm{d}r \\
&= \int_0^\infty \left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{rt^2}{2\nu}\right\} c(\nu) r^{\nu/2-1} e^{-r/2}\sqrt{\frac{r}{\nu}}\,\mathrm{d}r \\
&= \left(\frac{1}{2\pi}\right)^{1/2} \frac{c(\nu)}{\sqrt{\nu}} \int_0^\infty r^{(\nu+1)/2-1} \exp\left\{-\frac{r}{2}\left(1+\frac{t^2}{\nu}\right)\right\}\,\mathrm{d}r \\
&= \left(\frac{1}{2\pi}\right)^{1/2} \frac{c(\nu)}{\sqrt{\nu}} \left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \int_0^\infty z^{(\nu+1)/2-1} \exp\left\{-\frac{z}{2}\right\}\,\mathrm{d}z \quad \textit{setting } z = r\left(1+\frac{t^2}{\nu}\right) \\
&= \left(\frac{1}{2\pi}\right)^{1/2} \frac{c(\nu)}{\sqrt{\nu}} \left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \frac{1}{c(\nu+1)},
\end{aligned}$$

as the integrand is proportional to a Gamma pdf. We also can see that f_S is a $\text{Gamma}(\nu/2, 1/2)$ *(otherwise known as a $\chi^2(\nu)$ or χ^2_ν) density, and that the normalizing constant $c(\nu)$ is given by*

$$c(\nu) = \frac{\left(\frac{1}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} \qquad \Longrightarrow \qquad f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{1}{\pi\nu}\right)^{1/2} \frac{1}{(1+t^2/\nu)^{(\nu+1)/2}},$$

which, in fact, is the density of the $\text{Student}(\nu)$ *or t_ν distribution. You will recall this from your work on the t-test last year.*

For discussion

6. 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 $f(x_1|z = 0)$.

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

   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

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

   (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 tranformation methods,

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

   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

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

   (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 $\lim_{n\to\infty} \Pr(A_n) = 0$. When for W we approach the limit using a different sequence, unsurprisingly, we get a different answer.

7. Consider the data in Table 1, taken from Richard Doll's 1950s study of smoking. The table shows per capita consumption of cigarettes in 11 countries in 1930, and the death rates from lung cancer for men in 1950.

   (a) Produce a scatter plot of the data.

   (b) Why does the study compare cigarette consumption in 1930 with lung cancer rates 20 years later?

   (c) Why does the study only consider death rates in men?

   (d) Is it fair to conclude from these data that, on the whole, the higher the rate of smoking in a country in 1930, the higher the death rate from lung cancer in 1950?

   (e) Is it fair to conclude from these data that lung cancer death rates amongst smokers tend to be higher?

Table 1: Data on smoking and lung cancer rates

Country	Cigarette consumption	Deaths per million
Australia	480	180
Canada	500	150
Denmark	380	170
Finland	1100	350
Great Britain	1100	460
Iceland	230	60
Netherlands	490	240
Norway	250	90
Sweden	300	110
Switzerland	510	250
USA	1300	200

(a) See Figure 2.

(b) Lung cancer (supposing it results from smoking) is the cumulative effect of many years of exposure to cigarettes.

(c) In 1930, smoking was very uncommon amongst women.

(d) There is a positive correlation between the two variables($r \approx 0.74$), supporting the interpretation given.

(e) Not a reasonable conclusion from the data (although it is true). What we know is that the countries consuming more cigarettes tend to have higher incidence of lung cancer - these data do not say whether or not it is the people who consume the cigarettes that are dying.

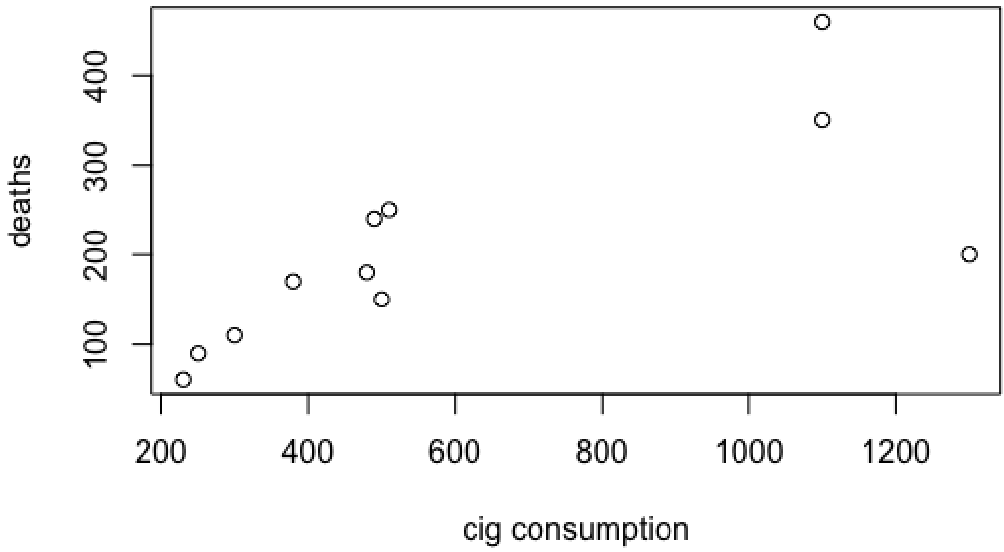

Figure 2: Scatter plot of cigarette consumption against lung cancer death rates.