MATH 556: MATHEMATICAL STATISTICS I

MULTIVARIATE DISTRIBUTION CALCULATIONS

Example 1: Let X_1 and X_2 be discrete rvs each with range $\{1, 2, 3, \ldots\}$ and joint mass function

$$f_{X_1,X_2}(x_1, x_2) = \frac{c}{(x_1 + x_2 - 1)(x_1 + x_2)(x_1 + x_2 + 1)} \qquad x_1, x_2 = 1, 2, 3, \ldots$$

and zero otherwise. The marginal mass function for X is given by

$$\begin{aligned} f_{X_1}(x_1) = \sum_{x_2=-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) &= \sum_{x_2=1}^{\infty} \frac{c}{(x_1 + x_2 - 1)(x_1 + x_2)(x_1 + x_2 + 1)} \\ &= \sum_{x_2=1}^{\infty} \frac{c}{2} \left[\frac{1}{(x_1 + x_2 - 1)(x_1 + x_2)} - \frac{1}{(x_1 + x_2)(x_1 + x_2 + 1)} \right] \\ &= \frac{c}{2} \frac{1}{x_1(x_1 + 1)} \end{aligned}$$

as all other terms cancel, and to calculate c, note that

$$\sum_{x_1=-\infty}^{\infty} f_{X_1}(x_1) = \sum_{x_1=1}^{\infty} \frac{c}{2} \frac{1}{x_1(x_1 + 1)} = \frac{c}{2} \sum_{x_1=1}^{\infty} \left[\frac{1}{x_1} - \frac{1}{x_1 + 1} \right] = \frac{c}{2}$$

as all terms in the sum except the first cancel. Hence $c = 2$. Also, as the joint function is symmetric in form for X_1 and X_2, f_{X_1} and f_{X_2} are identical.

Example 2: Let X_1 and X_2 be continuous rvs with supports $\mathbb{X}_1 = \mathbb{X}_2 = (0, 1)$ and joint pdf defined by

$$f_{X_1,X_2}(x_1, x_2) = 4x_1x_2 \qquad 0 < x_1 < 1,\ 0 < x_2 < 1$$

and zero otherwise. For $0 < x_1, x_2 < 1$,

$$\begin{aligned} F_{X_1,X_2}(x_1, x_2) &= \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{X_1,X_2}(t_1, t_2)\, dt_1 dt_2 = \int_0^{x_2} \int_0^{x_1} 4t_1t_2\, dt_1 dt_2 \\ &= \left\{ \int_0^{x_1} 2t_1\, dt_1 \right\} \left\{ \int_0^{x_2} 2t_2\, dt_2 \right\} = (x_1x_2)^2 \end{aligned}$$

and a full specification for F_{X_1,X_2} is

$$F_{X_1,X_2}(x_1, x_2) = \begin{cases} 0 & x_1, x_2 \leq 0 \\ (x_1x_2)^2 & 0 < x_1, x_2 < 1 \\ x_1^2 & 0 < x_1 < 1, x_2 \geq 1 \\ x_2^2 & 0 < x_2 < 1, x_1 \geq 1 \\ 1 & x_1, x_2 \geq 1 \end{cases}$$

To calculate, for $c \in \mathbb{R}$,

$$P_{X_1,X_2}\left[\frac{X_1 + X_2}{2} < c \right]$$

we need to integrate f_{X_1,X_2} over the set $A_c = \{(x_1, x_2) : 0 < x_1, x_2 < 1, (x_1 + x_2)/2 < c\}$, that is, if $c = 1/2$,

$$P_{X_1,X_2}[(X_1 + X_2) < 1] = \int_0^1 \int_0^{1-x_1} 4x_1x_2\, dx_2 dx_1 = \int_0^1 2x_1(1 - x_1)^2\, dx_1 = \frac{1}{6}$$

Example 3: Let X_1, X_2 be continuous random variables with supports $\mathbb{X}_1 \equiv \mathbb{X}_2 \equiv [0, 1]$, and joint pdf

$$f_{X_1,X_2}(x_1, x_2) = 1 \qquad 0 \le x_1, x_2 \le 1$$

and zero otherwise. Let $Y = X_1 + X_2$. Then $\mathbb{Y} \equiv [0, 2]$,

$$F_Y(y) = P_Y[Y \le y] = P_{X_1,X_2}[X_1 + X_2 \le y]$$

To calculate $P[X_1 + X_2 \le y]$, need to integrate f_{X_1,X_2} over the set

$$A_y = \{(x_1, x_2) : 0 < x_1, x_2 < 1, x_1 + x_2 \le y\}$$

This region is a portion of the unit square (that is, $\mathbb{X}_1 \times \mathbb{X}_2$) ; the line $x_1 + x_2 = y$ is a line with negative slope that cuts the horizontal axis at $x_1 = y$, and the vertical axis at $x_2 = y$.

- For $0 \le y \le 1$, A_y is the dark shaded lower triangle in the left panel of the figure below; hence for fixed y,

$$P_{X_1,X_2}[X_1 + X_2 < y] = \int_0^y \int_0^{y-x_2} 1 \; dx_1 dx_2 = \int_0^y (y - x_2) dx_2 = \frac{y^2}{2}.$$

- For $1 \le y \le 2$, A_y is more complicated see the figure below (right panel). It is easier mathematically to describe the complement of A_y within $\mathbb{X}_1 \times \mathbb{X}_2$ (striped in the right panel of the figure below), so we instead compute the complement probability as follows:

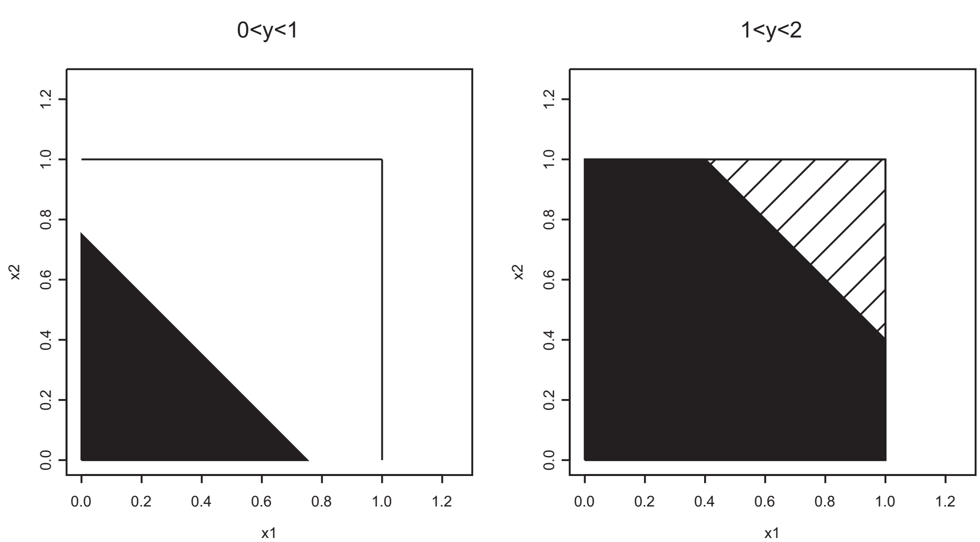
$$P_{X_1,X_2}[X_1 + X_2 \le y] = 1 - P_{X_1,X_2}[X_1 + X_2 > y]$$

$$= 1 - \int_{y-1}^1 \int_{y-x_2}^1 1 \; dx_1 dx_2 = 1 - \int_{y-1}^1 (1 - y + x_2) dx_2 = -\frac{y^2}{2} + 2y - 1$$

These two expressions give the cdf F_Y, and hence by differentiation we have

$$f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 2 - y & 1 \le y \le 2 \end{cases}$$

and zero otherwise.



Example 4: Let X_1 and X_2 be continuous rvs with supports $\mathbb{X}_1 = (0,1)$, $\mathbb{X}_2 = (0,2)$ and joint pdf

$$f_{X_1,X_2}(x_1,x_2) = c\left(x_1^2 + \frac{x_1 x_2}{2}\right) \qquad 0 < x_1 < 1,\ 0 < x_2 < 2$$

and zero otherwise.

(i) To calculate c, we have

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_1 dx_2 &= \int_0^2 \left\{\int_0^1 c\left(x_1^2 + \frac{x_1 x_2}{2}\right)\, dx_1\right\} dx_2 \\
&= \int_0^2 c\left[\frac{x_1^3}{3} + \frac{x_1^2 x_2}{4}\right]_0^1 dx_2 \\
&= \int_0^2 c\left(\frac{1}{3} + \frac{x_2}{4}\right)\, dx_2 \\
&= c\left[\frac{x_2}{3} + \frac{x_2^2}{8}\right]_0^2 = c\frac{7}{6}
\end{aligned}$$

so $c = 6/7$. The marginal pdf of X_1 is given, for $0 < x_1 < 1$, by

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_2 = \int_0^2 \frac{6}{7}\left(x_1^2 + \frac{x_1 x_2}{2}\right)\, dx_2 = \frac{6}{7}\left[x_1^2 x_2 + \frac{x_1 x_2^2}{4}\right]_0^2 = \frac{6x_1(2x_1+1)}{7}$$

and is zero otherwise.

(ii) To compute $P_{X_1,X_2}[X_1 > X_2]$, let

$$A = \{\ (x_1,x_2)\ :\ 0 < x_1 < 1, 0 < x_2 < 2, x_2 < x_1\}$$

so that

$$\begin{aligned}
P_{X_1,X_2}[X_1 &> X_2] = \iint\limits_A f_{X_1,X_2}(x_1,x_2)\, dx_2 dx_1 \\
&= \int_0^1 \left\{\int_0^{x_1} \frac{6}{7}\left(x_1^2 + \frac{x_1 x_2}{2}\right)\, dx_2\right\} dx_1 \\
&= \int_0^1 \left[x_1^2 x_2 + \frac{x_1 x_2^2}{4}\right]_0^{x_1} dx_1 \\
&= \int_0^1 \left(x_1^3 + \frac{x_1^3}{4}\right)\, dx_1 \\
&= \frac{6}{7}\left[\frac{5x_1^4}{16}\right]_0^1 \\
&= \frac{15}{56}
\end{aligned}$$

Example 5: Let X_1, X_2 and X_3 be continuous rvs with joint pdf defined by

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = c \qquad 0 < x_1 < x_2 < x_3 < 1$$

and zero otherwise. The support of this pdf is $\mathbb{X}^{(3)} = \{(x_1,x_2,x_3) : 0 < x_1 < x_2 < x_3 < 1\}$.

(i) To calculate c, integrate carefully over $\mathbb{X}^{(3)}$, that is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2,X_3}(x_1,x_2,x_3)\,dx_1\,dx_2\,dx_3 = 1$$

gives that

$$\int_0^1 \left\{\int_0^{x_3}\left\{\int_0^{x_2} c\,dx_1\right\}\,dx_2\right\}\,dx_3 = 1$$

Now

$$\int_0^1 \left\{\int_0^{x_3}\left\{\int_0^{x_2} c\,dx_1\right\}\,dx_2\right\}\,dx_3 = \int_0^1 \left\{\int_0^{x_3} cx_2\,dx_2\right\}\,dx_3 = \int_0^1 \frac{cx_3^2}{2}\,dx_3 = \frac{c}{6}$$

and hence $c = 6$.

(ii) For $0 < x_3 < 1$, f_{X_3} is given by

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2,X_3}(x_1,x_2,x_3)\,dx_1\,dx_2 = \int_0^{x_3}\left\{\int_0^{x_2} 6\,dx_1\right\}\,dx_2 = \int_0^{x_3} 6x_2\,dx_2 = 3x_3^2$$

and is zero otherwise. Similar calculations for X_1 and X_2 give

$$\begin{aligned} f_{X_1}(x_1) &= 3(1-x_1)^2 \qquad 0 < x_1 < 1 \\ f_{X_2}(x_2) &= 6x_2(1-x_2) \qquad 0 < x_2 < 1 \end{aligned}$$

with both densities equal to zero outside of these supports. Furthermore, for the **joint marginal** of X_1 and X_2, we have

$$f_{X_1,X_2}(x_1,x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2,X_3}(x_1,x_2,x_3)\,dx_3 = \int_{x_2}^{1} 6\,dx_3 = 6(1-x_2) \qquad 0 < x_1 < x_2 < 1$$

and zero otherwise. We have for the conditional of X_1 given $X_2 = x_2$,

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)} = \frac{1}{x_2} \qquad 0 < x_1 < x_2$$

and zero otherwise for **fixed** x_2.

(iii) We can calculate the expectation of X_1 directly

$$\mathbb{E}_{X_1}[X_1] = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)\,dx_1 = \int_0^1 x_1\,3(1-x_1)^2\,dx_1 = \frac{1}{4}$$

or, alternatively, using the *law of iterated expectation* (see page 11)

$$\mathbb{E}_{X_1|X_2}[X_1|X_2 = x_2] = \int_{-\infty}^{\infty} x_1 f_{X_1|X_2}(x_1|x_2)\,dx_1 = \int_0^{x_2} x_1\frac{1}{x_2}\,dx_1 = \frac{x_2}{2}$$

and hence by the law of iterated expectation

$$\begin{aligned} \mathbb{E}_{X_1}[X_1] &= \mathbb{E}_{X_2}\left[\mathbb{E}_{X_1|X_2}[X_1|X_2]\right] = \int_{-\infty}^{\infty}\left\{\mathbb{E}_{X_1|X_2}[X_1|X_2 = x_2]\right\} f_{X_2}(x_2)dx_2 \\ &= \int_0^1 \frac{x_2}{2} 6x_2(1-x_2)dx_2 = \frac{1}{4} \end{aligned}$$

Multivariate 1-1 Transformations

We consider the case of 1-1 transformations g, as in this case the probability transform result coincides with changing variables in a d-dimensional integral. We can consider $g = (g_1, \dots, g_d)$ as a vector of functions forming the components of the new random vector $\mathbf{Y}$.

Given a collection of variables $(X_1, \dots, X_d)$ with support $\mathbb{X}^{(d)}$ and joint pdf $f_{X_1,\dots,X_d}$ we can construct the pdf of a transformed set of variables $(Y_1, \dots, Y_d)$ using the following steps:

(I) Write down the set of transformation functions $g_1, \dots, g_d$

$$\begin{aligned} Y_1 &= g_1(X_1, \dots, X_d) \\ &\vdots \\ Y_d &= g_d(X_1, \dots, X_d) \end{aligned}.$$

(II) Write down the set of inverse transformation functions $g_1^{-1}, \dots, g_d^{-1}$

$$\begin{aligned} X_1 &= g_1^{-1}(Y_1, \dots, Y_d) \\ &\vdots \\ X_d &= g_d^{-1}(Y_1, \dots, Y_d) \end{aligned}$$

(III) Consider the joint support of the new variables, $\mathbb{Y}^{(d)}$.

(IV) Compute the Jacobian of the transformation: first form the matrix of partial derivatives

$$D_y = \begin{bmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} & \cdots & \dfrac{\partial x_1}{\partial y_d} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} & \cdots & \dfrac{\partial x_2}{\partial y_d} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_d}{\partial y_1} & \dfrac{\partial x_d}{\partial y_2} & \cdots & \dfrac{\partial x_d}{\partial y_d} \end{bmatrix}$$

where, for each (i, j)

$$\frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j}\left\{g_i^{-1}(y_1, \dots, y_d)\right\}$$

and then set $|J(y_1, \dots, y_d)| = |\det D_y|$

Note that

$$\det D_y = \det D_y^\top$$

so that an alternative but equivalent Jacobian calculation can be carried out by forming $D_y^\top$. *Note also that*

$$|J(y_1, \dots, y_d)| = \frac{1}{|J(x_1, \dots, x_d)|}$$

where $J(x_1, \dots, x_d)$ *is the Jacobian of the transformation regarded in the reverse direction (that is, if we start with* $(Y_1, \dots, Y_d)$ *and transfrom to* $(X_1, \dots, X_d)$*)*

(V) Write down the joint pdf of $(Y_1, \dots, Y_d)$ as

$$f_{Y_1,\dots,Y_d}(y_1, \dots, y_d) = f_{X_1,\dots,X_d}\left(g_1^{-1}(y_1, \dots, y_d), \dots, g_d^{-1}(y_1, \dots, y_d)\right) \times |J(y_1, \dots, y_d)|$$

for $(y_1, \dots, y_d) \in \mathbb{Y}^{(d)}$

Example 6: Suppose that X_1 and X_2 have joint pdf

$$f_{X_1,X_2}(x_1,x_2) = 2 \qquad 0 < x_1 < x_2 < 1$$

and zero otherwise. Compute the joint pdf of random variables

$$Y_1 = \frac{X_1}{X_2} \qquad\qquad Y_2 = X_2$$

SOLUTION

(I) Given that $\mathbb{X}^{(2)} \equiv \{(x_1,x_2) : 0 < x_1 < x_2 < 1\}$ and

$$g_1(t_1,t_2) = \frac{t_1}{t_2} \qquad\qquad g_2(t_1,t_2) = t_2$$

(II) Inverse transformations:

$$\left.\begin{array}{l} Y_1 = X_1/X_2 \\ Y_2 = X_2 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{l} X_1 = Y_1Y_2 \\ X_2 = Y_2 \end{array}\right.$$

and thus

$$g_1^{-1}(t_1,t_2) = t_1t_2 \qquad\qquad g_2^{-1}(t_1,t_2) = t_2$$

(III) Range: to find $\mathbb{Y}^{(2)}$ consider point by point transformation from $\mathbb{X}^{(2)}$ to $\mathbb{Y}^{(2)}$ For a pair of points $(x_1,x_2) \in \mathbb{X}^{(2)}$ and $(y_1,y_2) \in \mathbb{Y}^{(2)}$ linked via the transformation, we have

$$0 < x_1 < x_2 < 1 \Longleftrightarrow 0 < y_1y_2 < y_2 < 1$$

and hence we can extract the inequalities

$$0 < y_2 < 1 \text{ and } 0 < y_1 < 1 \qquad \therefore \qquad \mathbb{Y}^{(2)} \equiv (0,1) \times (0,1)$$

(IV) The Jacobian for points $(y_1,y_2) \in \mathbb{Y}^{(2)}$ is

$$D_y = \begin{bmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{bmatrix} = \begin{bmatrix} y_2 & y_1 \\ 0 & 1 \end{bmatrix} \Rightarrow |J(y_1,y_2)| = |\det D_y| = |y_2| = y_2$$

Note that for points $(x_1,x_2) \in \mathbb{X}^{(2)}$ is

$$D_x = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{x_2} & \dfrac{x_1}{x_2^2} \\ 0 & 1 \end{bmatrix} \Rightarrow |J(x_1,x_2)| = |\det D_x| = \left|\frac{1}{x_2}\right| = \frac{1}{x_2}$$

so that

$$|J(y_1,y_2)| = \frac{1}{|J(x_1,x_2)|}$$

(V) Finally, we have

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(y_1y_2,y_2) \times y_2 = 2y_2 \qquad 0 < y_1 < 1, 0 < y_2 < 1$$

and zero otherwise

Example 7: Suppose that X_1 and X_2 are **independent** and **identically distributed** random variables defined on $\mathbb{R}^+$ each with pdf of the form

$$f_X(x) = \sqrt{\frac{1}{2\pi x}} \exp\left\{-\frac{x}{2}\right\} \qquad x > 0$$

and zero otherwise. Compute the joint pdf of random variables $Y_1 = X_1$ and $Y_2 = X_1 + X_2$

SOLUTION

(I) Given that $\mathbb{X}^{(2)} \equiv \{(x_1, x_2) : 0 < x_1, 0 < x_2\}$ and

$$g_1(t_1, t_2) = t_1 \qquad\qquad g_2(t_1, t_2) = t_1 + t_2$$

(II) Inverse transformations:

$$\left.\begin{array}{l} Y_1 = X_1 \\ Y_2 = X_1 + X_2 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{l} X_1 = Y_1 \\ X_2 = Y_2 - Y_1 \end{array}\right.$$

and thus

$$g_1^{-1}(t_1, t_2) = t_1 \qquad\qquad g_2^{-1}(t_1, t_2) = t_2 - t_1$$

(III) Range: to find $\mathbb{Y}^{(2)}$ consider point by point transformation from $\mathbb{X}^{(2)}$ to $\mathbb{Y}^{(2)}$ For a pair of points $(x_1, x_2) \in \mathbb{X}^{(2)}$ and $(y_1, y_2) \in \mathbb{Y}^{(2)}$ linked via the transformation; as both original variables are strictly positive, we can extract the inequalities

$$0 < y_1 < y_2 < \infty$$

(IV) The Jacobian for points $(y_1, y_2) \in \mathbb{Y}^{(2)}$ is

$$D_y = \begin{bmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \Rightarrow |J(y_1, y_2)| = |\det D_y| = |1| = 1$$

Note, here, $J(x_1, x_2) = |\det D_x| = 1$ also so that again

$$|J(y_1, y_2)| = \frac{1}{|J(x_1, x_2)|}$$

(V) Finally, we have for $0 < y_1 < y_2 < \infty$

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}(y_1, y_2 - y_1) \times 1 = f_{X_1}(y_1) \times f_{X_2}(y_2 - y_1) \qquad \text{by independence} \\ &= \sqrt{\frac{1}{2\pi y_1}} \exp\left\{-\frac{y_1}{2}\right\} \sqrt{\frac{1}{2\pi(y_2 - y_1)}} \exp\left\{-\frac{(y_2 - y_1)}{2}\right\} \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{y_1(y_2 - y_1)}} \exp\left\{-\frac{y_2}{2}\right\} \end{aligned}$$

and zero otherwise

Here, for $y_2 > 0$

$$
\begin{aligned}
f_{Y_2}(y_2) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1, y_2)\, dy_1 = \int_0^{y_2} \frac{1}{2\pi} \frac{1}{\sqrt{y_1(y_2 - y_1)}} \exp\left\{-\frac{y_2}{2}\right\} dy_1 \\
&= \frac{1}{2\pi} \exp\left\{-\frac{y_2}{2}\right\} \int_0^{y_2} \frac{1}{\sqrt{y_1(y_2 - y_1)}}\, dy_1 \\
&= \frac{1}{2\pi} \exp\left\{-\frac{y_2}{2}\right\} \int_0^{1} \frac{1}{\sqrt{ty_2(y_2 - ty_2)}} y_2\, dt \qquad \text{setting } y_1 = ty_2 \\
&= \frac{1}{2\pi} \exp\left\{-\frac{y_2}{2}\right\} \int_0^{1} \frac{1}{\sqrt{t(1-t)}}\, dt \\
&= \frac{1}{2} \exp\left\{-\frac{y_2}{2}\right\}
\end{aligned}
$$

as

$$\int_0^1 \frac{1}{\sqrt{t(1-t)}}\, dt = \pi$$

either by direct calculation, or by recognizing the integrand as proportional to a $Beta(1/2, 1/2)$ pdf.

Example 8: The Cauchy distribution is a symmetric distribution on $(-\infty, \infty)$ with pdf

$$f_X(x; \theta, \sigma) = \frac{1}{\pi}\frac{1}{\sigma} \cdot \frac{1}{1 + \left(\dfrac{x-\theta}{\sigma}\right)^2} = \frac{1}{\pi} \cdot \frac{\sigma}{\sigma^2 + (x-\theta)^2}$$

The standard case is $\theta = 0$, $\sigma = 1$.

The Cauchy distribution arises as the ratio of two independent Gaussian random variables. Suppose that $X, Y \sim Normal(0, 1)$. We then proceed by

(a) defining the transformation $U = X/Y$ and $V = |Y|$,

(b) finding the joint pdf $f_{U,V}(u, v)$, and

(c) integrating out V to obtain the marginal pdf of U.

Overall, the mapping $U = X/Y$ and $V = |Y|$ is not 1-1: the two points (x, y) and $(-x, -y)$ map to the same (u, v). However, we may partition the support of (X, Y) into three regions A_0, A_1, A_2 such that the mapping from A_i to (U, V) is one-to-one on each. For simplicity here we denote the inverse mappings as h rather than g^{-1}.

(i) $A_0 = \{(X, Y) : Y = 0\}$: we can ignore this case as the distribution of Y is continuous, so $P_Y[Y = 0] = 0$ when $Y \sim Normal(0, 1)$.

(ii) $A_1 = \{(X, Y) : Y > 0\}$: The mapping $U = X/Y$, $V = |Y|$ is 1-1, and the inverse mappings are $h_{11}(u, v) = uv$, $h_{21}(u, v) = v$.

(iii) $A_2 = \{(X, Y) : Y < 0\}$: The mapping $U = X/Y$, $V = |Y|$ is one-to-one, and the inverse mappings are $h_{12}(u, v) = -uv$, $h_{22}(u, v) = -v$.

In cases (ii) and (iii) we have the following Jacobians:

$$J_1 = \begin{vmatrix} \dfrac{\partial h_{11}(u,v)}{\partial u} & \dfrac{\partial h_{11}(u,v)}{\partial v} \\ \dfrac{\partial h_{21}(u,v)}{\partial u} & \dfrac{\partial h_{21}(u,v)}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial (uv)}{\partial u} & \dfrac{\partial (uv)}{\partial v} \\ \dfrac{\partial v}{\partial u} & \dfrac{\partial v}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$$

$$J_2 = \begin{vmatrix} \dfrac{\partial h_{12}(u,v)}{\partial u} & \dfrac{\partial h_{12}(u,v)}{\partial v} \\ \dfrac{\partial h_{22}(u,v)}{\partial u} & \dfrac{\partial h_{22}(u,v)}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial (-uv)}{\partial u} & \dfrac{\partial (-uv)}{\partial v} \\ \dfrac{\partial (-v)}{\partial u} & \dfrac{\partial (-v)}{\partial v} \end{vmatrix} = \begin{vmatrix} -v & -u \\ 0 & -1 \end{vmatrix} = v$$

We have that

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\} \frac{1}{\sqrt{2\pi}} \exp\{-y^2/2\} = \frac{1}{2\pi} \exp\left\{-\frac{(x^2+y^2)}{2}\right\}$$

so therefore, using the indicator function to delineate the two cases, we have

$$\begin{aligned} f_{U,V}(u,v) &= \mathbb{1}_{A_1}(u,v) f_{X,Y}(h_{11}(u,v), h_{21}(u,v))|J_1| + \mathbb{1}_{A_2}(u,v) f_{X,Y}(h_{12}(u,v), h_{22}(u,v))|J_2| \\ &= \frac{\mathbb{1}_{A_1}(u,v)}{2\pi} \exp\left(-\frac{(uv)^2+v^2}{2}\right)|v| + \frac{\mathbb{1}_{A_2}(u,v)}{2\pi} \exp\left(-\frac{(-uv)^2+(-v)^2}{2}\right)|v| \\ &= \frac{v}{\pi} \exp\left(-\frac{v^2(u^2+1)}{2}\right), \qquad u \in \mathbb{R}, v \in \mathbb{R}^+ \end{aligned}$$

and hence, on marginalization

$$\begin{aligned} f_U(u) &= \int_0^\infty \frac{v}{\pi} \exp\left\{-\frac{v^2(u^2+1)}{2}\right\} dv & &\text{integrating out } v \\ &= \int_0^\infty \frac{1}{2\pi} \exp\left\{-\frac{(u^2+1)}{2} z\right\} dz & &\text{setting } z = v^2 \text{ and } dz = 2v dv \\ &= \frac{1}{2\pi} \cdot \frac{2}{1+u^2} & &\int_0^\infty \exp(-\alpha z) dz = \frac{1}{\alpha} \\ &= \frac{1}{\pi} \cdot \frac{1}{1+u^2} \end{aligned}$$

The general $Cauchy(\theta, \sigma)$ form is generated using a linear transformation: if $Z \sim Cauchy(0,1)$, then

$$X = \sigma Z + \theta$$

has a $Cauchy(\theta, \sigma)$ distribution. The second (equivalent) construction of the standard Cauchy distribution is as a *scale mixture*. Suppose X and Y have a joint distribution specified as

$$\begin{aligned} Y &\sim \chi^2_1 \equiv Gamma(1/2, 1/2) \\ X|Y = y &\sim Normal(0, y^{-1}) \end{aligned}$$

that is, the variance of X given $Y = y$ is $1/y$. Then we have that

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy = \int_0^{\infty} f_{X|Y}(x|y) f_Y(y)\, dy$$

$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi}} y^{1/2} \exp\left\{-\frac{y}{2}x^2\right\} \frac{(1/2)^{1/2}}{\Gamma(1/2)} y^{-1/2} \exp\left\{-\frac{y}{2}\right\}\, dy$$

$$= \frac{1}{2\pi} \int_0^{\infty} \exp\left\{-\frac{y}{2}(1+x^2)\right\}\, dy$$

$$= \frac{1}{\pi} \frac{1}{1+x^2}$$

as $\Gamma(1/2) = \sqrt{\pi}$.

Example 9: Let X_1, X_2 be continuous random variables with joint density f_{X_1,X_2} and let rv Y be defined by $Y = g(X_1, X_2)$. To calculate the pdf of Y we could use the multivariate transformation theorem after defining another (dummy) variable Z as some function of X_1 and X_2, and consider the joint transformation $(X_1, X_2) \longrightarrow (Y, Z)$. Defining $Z = X_1$, we have

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y,Z}(y,z)\, dz = \int_{-\infty}^{\infty} f_{Y|Z}(y|z) f_Z(z)\, dz = \int_{-\infty}^{\infty} f_{Y|X_1}(y|x_1) f_{X_1}(x_1)\, dx_1$$

as $f_{Y,Z}(y,z) = f_{Y|Z}(y|z) f_Z(z)$ by the chain rule for densities; $f_{Y|X_1}(y|x_1)$ is a univariate (conditional) pdf for Y given $X_1 = x_1$.

Now, **given** that $X_1 = x_1$, we have that $Y = g(x_1, X_2)$, that is, Y is a transformation of X_2 only. Hence the conditional pdf $f_{Y|X_1}(y|x_1)$ can be derived using single variable (rather than multivariate) transformation techniques. Specifically, if $Y = g(x_1, X_2)$ is a 1-1 transformation from X_2 to Y, then the inverse transformation $X_2 = g^{-1}(x_1, Y)$ is well defined, and by the transformation theorem

$$f_{Y|X_1}(y|x_1) = f_{X_2|X_1}(g^{-1}(x_1,y))\, |J(y;x_1)| = f_{X_2|X_1}(g^{-1}(x_1,y)|x_1) \left| \frac{\partial}{\partial t} \left\{g^{-1}(x_1,t)\right\}_{t=y} \right|$$

and hence

$$f_Y(y) = \int_{-\infty}^{\infty} \left\{ f_{X_2|X_1}(g^{-1}(x_1,y)|x_1) \left| \frac{\partial}{\partial t} \left\{g^{-1}(x_1,t)\right\}_{t=y} \right| \right\} f_{X_1}(x_1) dx_1$$

For example, if $Y = X_1 X_2$, then $X_2 = Y/X_1$, and hence

$$\left| \frac{\partial}{\partial t} \left\{g^{-1}(x_1,t)\right\}_{t=y} \right| = \left| \frac{\partial}{\partial t} \left\{ \frac{t}{x_1} \right\}_{t=y} \right| = |x_1|^{-1}$$

so

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2|X_1}(y/x_1|x_1)\, |x_1|^{-1}\, f_{X_1}(x_1) dx_1.$$

The conditional density $f_{X_2|X_1}$ and/or the marginal density f_{X_1} may be zero on parts of the range of the integral. Alternatively, the **cdf** of Y is given by

$$F_Y(y) = P[Y \le y] = P[g(X_1, X_2) \le y] = \iint_{A_y} f_{X_1,X_2}(x_1,x_2)\, dx_2 dx_1$$

where $A_y = \{\ (x_1,x_2)\ : g(x_1,x_2) \le y\}$ so the cdf can be calculated by carefully identifying and intergrating over the set A_y.

Multivariate Expectations

We define a multivariate expectation using the same approach as in the univariate case. If $X = (X_1, \ldots, X_d$ is a d-dimensional random vector, and g is a k-dimensional function, then

$$\mathbb{E}_X[g(X)] = \int g(x)\, dF_X(x)$$

that is, in the discrete case

$$\mathbb{E}_{X_1,\ldots,X_d}[g(X_1,\ldots,X_d)] = \int_{x\in\mathbb{R}^d} g(x_1,\ldots,x_d) f_{X_1,\ldots,X_d}(x_1,\ldots,x_d)$$

and in the continuous case

$$\mathbb{E}_{X_1,\ldots,X_d}[g(X_1,\ldots,X_d)] = \sum_{x\in\,\mathbb{X}} g(x_1,\ldots,x_d) f_{X_1,\ldots,X_d}(x_1,\ldots,x_d)\, dx_1 \ldots dx_d$$

Example 10: The **law of iterated expectation** uses a decomposition of the joint pmf or pdf to compute an expectation. For example, let X_1, X_2 be rvs with joint density f_{X_1,X_2}. Then

$$\begin{aligned}
\mathbb{E}_{X_1}[X_1] &= \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)\, dx_1 \\
&= \int_{-\infty}^{\infty} x_1 \left\{ \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_2 \right\} dx_1 && \text{defn of marginal} \\
&= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x_1 f_{X_1|X_2}(x_1|x_2) f_{X_2}(x_2) dx_2 \right\} dx_2 && \text{exch. order of intgn.} \\
&= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x_1 f_{X_1|X_2}(x_1|x_2)\, dx_1 \right\} f_{X_2}(x_2) dx_2 \\
&= \mathbb{E}_{X_2}\left[\mathbb{E}_{X_1|X_2}[X_1|X_2]\right]
\end{aligned}$$

as the inner integral is the **conditional expectation**

$$\mathbb{E}_{X_1|X_2}[X_1|X_2 = x_2] = \int_{-\infty}^{\infty} x_1 f_{X_1|X_2}(x_1|x_2)\, dx_1.$$

Let $g(X_1)$ be a function of X_1 only. Then

$$\begin{aligned}
\mathbb{E}_{X_1,X_2}[g(X_1)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1) f_{X_1,X_2}(x_1,x_2) dx_1 dx_2 \\
&= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(x_1) f_{X_1|X_2}(x_1|x_2) f_{X_2}(x_2) dx_1 \right\} dx_2 \\
&= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(x_1) f_{X_1|X_2}(x_1|x_2) dx_1 \right\} f_{X_2}(x_2) dx_2 \\
&= \mathbb{E}_{X_2}\left[\mathbb{E}_{X_1|X_2}[g(X_1)|X_2]\right] = \mathbb{E}_{X_1}[g(X_1)]
\end{aligned}$$

by the law of iterated expectation. Thus, we can compute the expectation with respect to the marginal f_{X_1} rather than the joint pdf.

Example 11: If X_1 and X_2 are continuous rvs with joint mass function/pdf f_{X_1,X_2}, then the **covariance** of X_1 and X_2 is defined by

$$\begin{aligned}\text{Cov}_{X_1,X_2}[X_1,X_2] &= \mathbb{E}_{X_1,X_2}[(X_1-\mu_1)(X_2-\mu_2)] \\ &= \iint (x_1-\mu_1)(x_2-\mu_2)\, f_{X_1,X_2}(x_1,x_2)\, dx_1 dx_2 \\ &= \mathbb{E}_{X_1,X_2}[X_1X_2] - \mu_2\mathbb{E}_{X_1}[X_1] - \mu_1\mathbb{E}_{X_2}[X_2] + \mu_1\mu_2 \\ &= \mathbb{E}_{X_1,X_2}[X_1X_2] - \mu_1\mu_2\end{aligned}$$

where $\mu_i = \mathbb{E}_{X_i}[X_i]$ is the marginal expectation of X_i, for $i = 1, 2$

It follows that if $Y = X_1 + X_2$, then

$$\begin{aligned}\mathbb{E}_Y[Y] &= \mathbb{E}_{X_1,X_2}[X_1+X_2] = \iint (x_1+x_2) f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 \\ &= \iint x_1 f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 + \iint x_2 f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 \\ &= \mathbb{E}_{X_1}[X_1] + \mathbb{E}_{X_2}[X_2]\end{aligned}$$

and

$$\begin{aligned}\text{Var}_Y[Y] &= \text{Var}_{X_1,X_2}[X_1+X_2] = \mathbb{E}_{X_1,X_2}\left[\, (X_1+X_2-(\mu_1+\mu_2))^2\right] \\ &= \iint (x_1+x_2-\mu_1-\mu_2)^2\, f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 \\ &= \iint \left[(x_1-\mu_1)^2 + (x_2-\mu_2)^2 + 2(x_1-\mu_1)(x_2-\mu_2)\right]\, f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 \\ &= \iint (x_1-\mu_1)^2 f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 + \iint (x_2-\mu_2)^2 f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 \\ &\quad +2\iint (x_1-\mu_1)(x_2-\mu_2)\, f_{X_1,X_2}(x_1,x_2)\, dx_1dx_2 \\ &= \text{Var}_{X_1}[X_1] + \text{Var}_{X_2}[X_2] + 2\,\text{Cov}_{X_1,X_2}[X_1,X_2]\end{aligned}$$

and the result for the sum of n variables follows similarly, or by induction.

Example 12: Let X_1, X_2 be continuous random variables with joint pdf given by

$$f_{X_1,X_2}(x_1,x_2) = c \qquad 0 < x_1 < 1, x_1 < x_2 < x_1+1$$

and zero otherwise. To calculate c, we have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_2dx_1 = \int_0^1\int_{x_1}^{x_1+1} c\, dx_2dx_1 = \int_0^1 c\,[x_2]_{x_1}^{x_1+1}\, dx_1 = \int_0^1 c\, dx_2 = c$$

so $c = 1$. The marginal pdf of X_1 is given by

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_2 = \int_{x_1}^{x_1+1} 1\, dx_2 = 1 \qquad 0 < x_1 < 1$$

and zero otherwise, and the marginal pdf for X_2 is given by

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\,dx_1 = \begin{cases} \displaystyle\int_0^{x_2} 1\,dx_1 & = x_2 & 0 < x_2 < 1 \\ \displaystyle\int_{x_2-1}^{1} 1\,dx_1 & = 2 - x_2 & 1 \le x_2 < 2 \end{cases}$$

and zero otherwise. Hence

$$\mathbb{E}_{X_1}[X_1] = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)\,dx_1 = \int_0^1 x_1\,dx_1 = \frac{1}{2}$$

$$\text{Var}_{X_1}[X_1] = \int_{-\infty}^{\infty} x_1^2 f_{X_1}(x_1)\,dx_1 - \{\mathbb{E}_{X_1}[X_1]\}^2 = \int_0^1 x_1^2\,dx_1 - \frac{1}{4} = \frac{1}{12}$$

$$\begin{aligned} \mathbb{E}_{X_2}[X_2] &= \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2)\,dx_2 = \int_0^1 x_2^2\,dx_2 + \int_1^2 x_2(2-x_2)\,dx_2 \\ &= \frac{1}{3} - \left(1 - \frac{1}{3}\right) + \left(4 - \frac{8}{3}\right) = 1 \end{aligned}$$

$$\begin{aligned} \text{Var}_{X_2}[X_2] &= \int_{-\infty}^{\infty} x_2^2 f_{X_2}(x_2)\,dx_2 - \{\mathbb{E}_{X_2}[X_2]\}^2 \\ &= \int_0^1 x_2^2 x_2\,dx_2 + \int_1^2 x_2^2(2-x_2)\,dx_2 - 1 \\ &= \frac{1}{4} - \left(\frac{2}{3} - \frac{1}{4}\right) + \left(\frac{16}{3} - 4\right) - 1 = \frac{1}{6} \end{aligned}$$

The covariance and correlation of X_1 and X_2 are then given by

$$\begin{aligned} \text{Cov}_{X_1,X_2}[X_1,X_2] &= \left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f_{X_1,X_2}(x_1,x_2)\,dx_2\right\} dx_1 - \mathbb{E}_{X_1}[X_1]\;\mathbb{E}_{X_2}[X_2] \\ &= \int_0^1 \left\{\int_{x_1}^{x_1+1} x_1x_2\,dx_2\right\} dx_1 - \frac{1}{2}.1 \\ &= \int_0^1 x_1\left[\frac{x_2}{2}\right]_{x_1}^{x_1+1} dx_1 \; - \; \frac{1}{2} \\ &= \int_0^1 \left(x_1^2 + \frac{x_1}{2}\right) dx_1 \; - \; \frac{1}{2} \\ &= \left[\frac{x_1^3}{3} + \frac{x_1^2}{4}\right]_0^1 - \frac{1}{2} = \frac{7}{12} - \frac{1}{2} = \frac{1}{12} \end{aligned}$$

and hence

$$\text{Corr}_{X_1,X_2}[X_1,X_2] = \frac{\text{Cov}_{X_1,X_2}[X_1,X_2]}{\sqrt{\text{Var}_{X_1}[X_1]\;\text{Var}_{X_2}[X_2]}} = \frac{1/12}{\sqrt{1/12}\sqrt{1/6}} = \frac{1}{\sqrt{2}}$$

Example 13: **Convolution Theorem** Suppose that X_1 and X_2 have a joint pmf or pdf, f_{X_1,X_2}, and let $Y = X_1 + X_2$. We compute the pmf/pdf of Y by using a Convolution Theorem, which for continuous variables is a special case of the transformation theorem.

- **Discrete Case:** By the Theorem of Total Probability, we have from first principles that for any fixed y.

$$f_Y(y) = P_Y[Y = y] = \sum_{x_1} \sum_{\substack{x_2 \\ x_1+x_2=y}} f_{X_1,X_2}(x_1, x_2) = \sum_{x_1} f_{X_1,X_2}(x_1, y - x_1)$$

- **Continuous Case:** Consider $Y = X_1 + X_2$ and $Z = X_1$. We have

$$\left. \begin{array}{l} Y = X_1 + X_2 \\ Z = X_1 \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{l} X_1 = Z \\ X_2 = Y - Z \end{array} \right.$$

The Jacobian of this transform is 1, so we conclude from the transformation result that for all (y, z)

$$f_{Y,Z}(y, z) = f_{X_1,X_2}(z, y - z)$$

and hence, marginalizing z, we see that

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y,Z}(y, z) \, dz = \int_{-\infty}^{\infty} f_{X_1,X_2}(z, y - z) \, dz$$

which we may rewrite

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, y - x_1) \, dx_1.$$

(i) We should establish explicitly the support of the new variable Y when recording f_Y.

(ii) The marginalization over x_1 must take into account the support of f_{X_1,X_2}: that is, for any fixed y only contributions to the sum or integral where

$$f_{X_1,X_2}(x_1, y - x_1) > 0.$$

Example 14: Let X_1, X_2 be continuous random variables with joint pdf given by

$$f_{X_1,X_2}(x_1, x_2) = x_1 \exp\{-(x_1 + x_2)\} \qquad x_1, x_2 > 0$$

and zero otherwise. Let $Y = X_1 + X_2$. Then by the Convolution Theorem, for $y > 0$,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, y - x_1) \, dx_1 \\ &= \int_0^y x_1 \exp\{-(x_1 + (y - x_1))\} \, dx_1 \qquad \text{as } f_{X_1,X_2}(x_1, y - x_1) > 0 \Leftrightarrow 0 < x_1 < y \\ &= \frac{1}{2} y^2 e^{-y} \qquad y > 0 \end{aligned}$$

and zero otherwise. Note that the integral range reduces to 0 to y as the joint density f_{X_1,X_2} is only non-zero when both its arguments are positive, that is, when $x_1 > 0$ and $y - x_1 > 0$ for fixed y, or when $0 < x_1 < y$. We conclude that $Y \sim Gamma(3, 1)$.

Example 15: Let X_1, X_2 be continuous random variables with joint pdf given by

$$f_{X_1,X_2}(x_1,x_2) = 2(x_1+x_2) \qquad 0 \le x_1 \le x_2 \le 1$$

and zero otherwise. Let $Y = X_1 + X_2$. Clearly Y takes values on $\mathbb{Y} \equiv [0,2]$.

For fixed y, $0 \le y \le 2$, we need to consider two ranges to respect the fact that the joint pdf is only non-zero if

$$0 \le x_1 \le x_2 \le 1$$

(i) For $0 \le y \le 1$:

$$0 \le x_1 \le y - x_1 \le 1 \quad \Longrightarrow \quad 0 \le 2x_1 \le y,$$

or equivalently $0 \le x_1 \le y/2$.

(ii) For $1 \le y \le 2$

$$0 \le x_1 \le y - x_1 \le 1 \quad \Longrightarrow \quad y-1 \le x_1 \le y/2.$$

Therefore, by the Convolution Theorem, as

$$f_{X_1,X_2}(x_1, y-x_1) = 2y$$

when the function is non-zero, we have

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, y-x_1)\,dx_1 = \begin{cases} \displaystyle\int_0^{y/2} 2y\,dx_1 & 0 \le y \le 1 \\ \displaystyle\int_{y-1}^{y/2} 2y\,dx_1 & 1 \le y \le 2 \end{cases}$$

and zero otherwise. Hence

$$f_Y(y) = \begin{cases} y^2 & 0 \le y \le 1 \\ y(2-y) & 1 \le y \le 2 \end{cases}$$

It is straightforward to check that this density is a valid pdf. The region of (X_1, Y) space on which the joint density $f_{X_1,X_2}(x_1, y-x_1)$ is **positive**; this region is the triangle with corners $(0,0), (1,2), (0,1)$.