

# MATH 323: PROBABILITY

## CALCULATIONS FOR MULTIVARIATE DISTRIBUTIONS

**EXAMPLE 1** Let  $Y_1$  and  $Y_2$  be discrete random variables each with range  $\{1, 2, 3, \dots\}$  and joint pmf

$$p_{Y_1, Y_2}(y_1, y_2) = \frac{c}{(y_1 + y_2 - 1)(y_1 + y_2)(y_1 + y_2 + 1)} \quad y_1, y_2 = 1, 2, 3, \dots$$

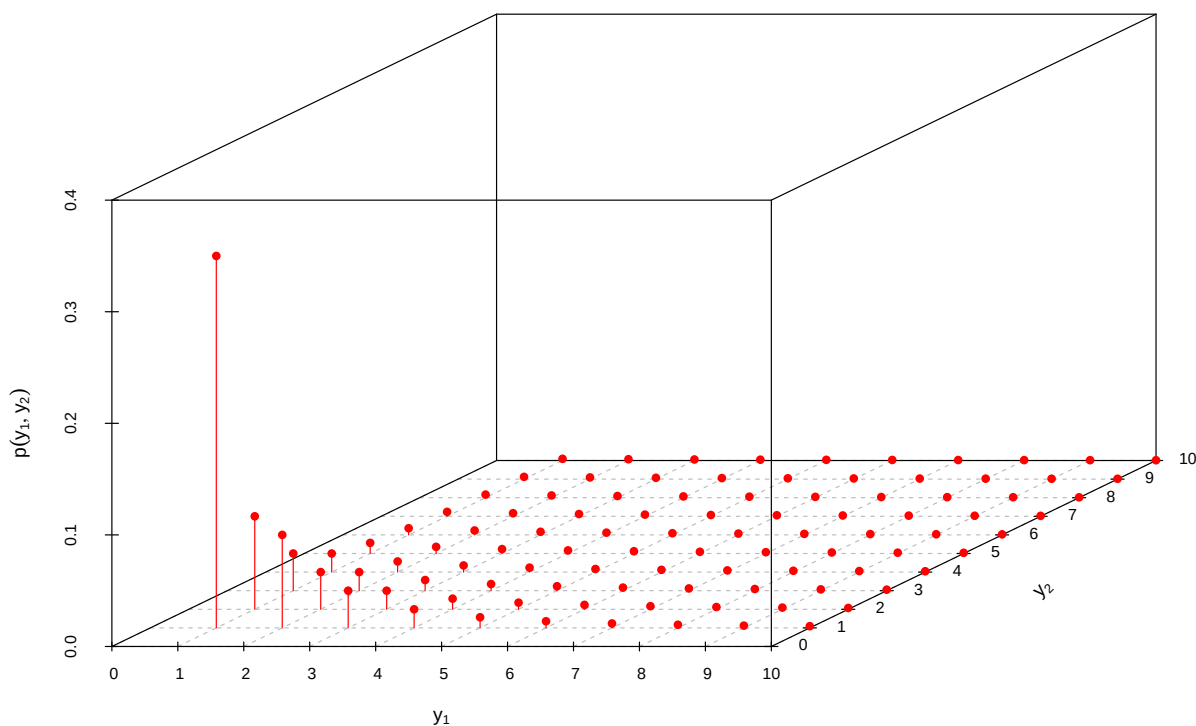
and zero otherwise. The marginal pmf for  $Y_1$  is given by

$$\begin{aligned} p_{Y_1}(y_1) &= \sum_{y_2=-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) = \sum_{y_2=1}^{\infty} \frac{c}{(y_1 + y_2 - 1)(y_1 + y_2)(y_1 + y_2 + 1)} \\ &= \sum_{y_2=1}^{\infty} \frac{c}{2} \left[ \frac{1}{(y_1 + y_2 - 1)(y_1 + y_2)} - \frac{1}{(y_1 + y_2)(y_1 + y_2 + 1)} \right] \\ &= \frac{c}{2} \frac{1}{y_1(y_1 + 1)} \end{aligned}$$

as all terms apart from the first cancel in the infinite sum. To calculate  $c$ , note that

$$\sum_{y_1=-\infty}^{\infty} p_{Y_1}(y_1) = \sum_{y_1=1}^{\infty} \frac{c}{2} \frac{1}{y_1(y_1 + 1)} = \frac{c}{2} \sum_{y_1=1}^{\infty} \left[ \frac{1}{y_1} - \frac{1}{y_1 + 1} \right] = \frac{c}{2}$$

as again all terms in the sum except the first cancel. Hence  $c = 2$ . Note that, as the joint function is symmetric in form for  $Y_1$  and  $Y_2$ ,  $p_{Y_1}$  and  $p_{Y_2}$  are identical.



**EXAMPLE 2** Let  $Y_1$  and  $Y_2$  be continuous random variables defined on  $\mathcal{Y}_1 = \mathcal{Y}_2 = (0, 1)$  and joint pdf

$$f_{Y_1, Y_2}(y_1, y_2) = 4y_1y_2 \quad 0 < y_1 < 1, 0 < y_2 < 1$$

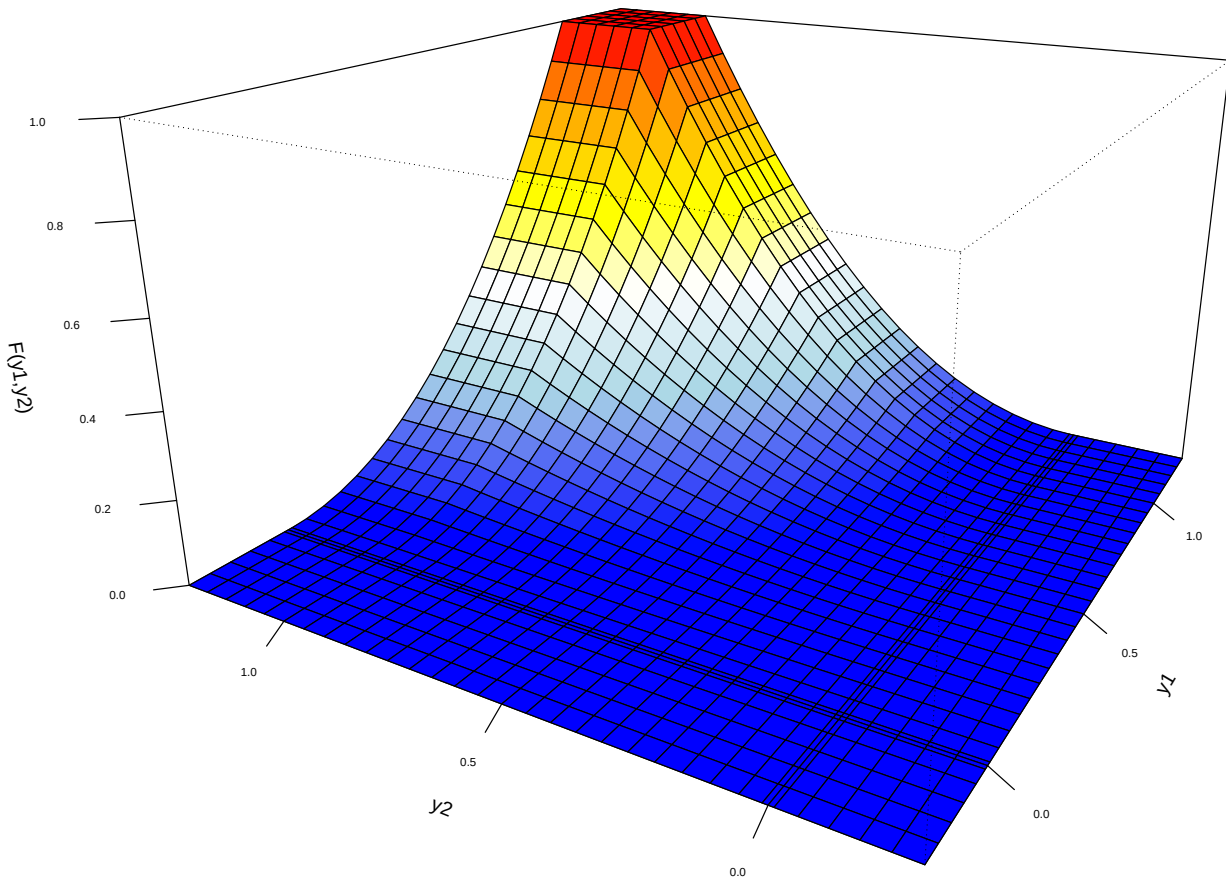
and zero otherwise. For  $0 < y_1, y_2 < 1$ ,

$$F_{Y_1, Y_2}(y_1, y_2) = \int_{-\infty}^{y_2} \int_{-\infty}^{y_1} f_{Y_1, Y_2}(t_1, t_2) dt_1 dt_2 = \int_0^{y_2} \int_0^{y_1} 4t_1 t_2 dt_1 dt_2 = \left\{ \int_0^{y_1} 2t_1 dt_1 \right\} \left\{ \int_0^{y_2} 2t_2 dt_2 \right\} = (y_1 y_2)^2$$

and a full specification for  $F_{Y_1, Y_2}$  is

$$F_{Y_1, Y_2}(y_1, y_2) = \begin{cases} 0 & y_1, y_2 \leq 0 \\ (y_1 y_2)^2 & 0 < y_1, y_2 < 1 \\ y_1^2 & 0 < y_1 < 1, y_2 \geq 1 \\ y_2^2 & 0 < y_2 < 1, y_1 \geq 1 \\ 1 & y_1, y_2 \geq 1 \end{cases}$$

The function is plotted as a surface plot below:



For fixed  $y$ , to calculate  $P(Y_1 + Y_2 < 1)$  we need to integrate  $f_{Y_1, Y_2}$  over the corresponding region, that is, set  $A = \{(y_1, y_2) : 0 < y_1, y_2 < 1, (y_1 + y_2) < 1\}$ , that is,

$$P(Y_1 + Y_2 < 1) = \int_0^1 \int_0^{1-y_1} 4y_1y_2 dy_2 dy_1 = \int_0^1 2y_1(1-y_1)^2 dy_1 = \frac{1}{6}$$

**EXAMPLE 3** Let  $Y_1, Y_2$  be continuous random variables with ranges  $\mathcal{Y}_1 \equiv \mathcal{Y}_2 \equiv [0, 1]$ , and joint pdf

$$f_{Y_1, Y_2}(y_1, y_2) = 1 \quad 0 \leq y_1, y_2 \leq 1$$

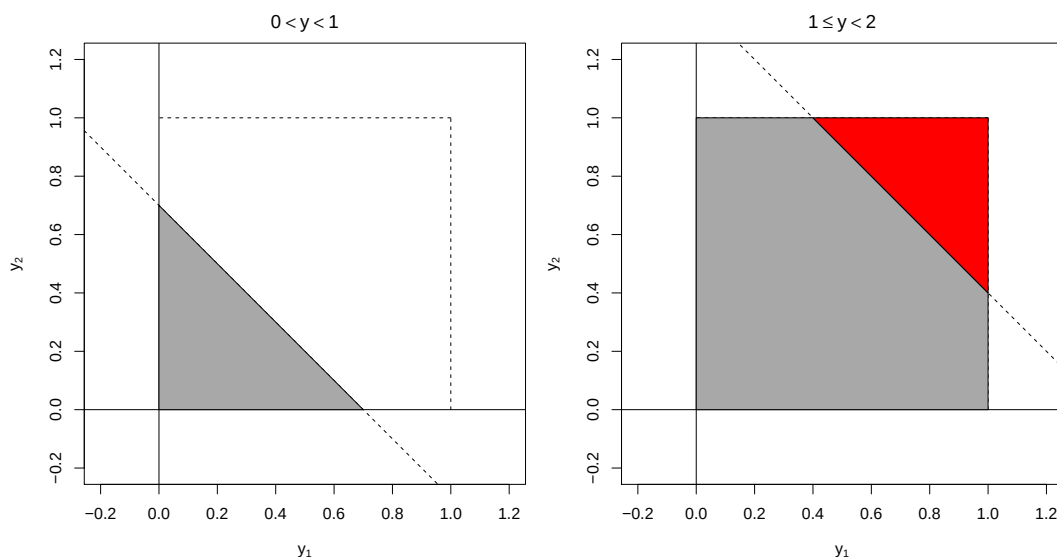
and zero otherwise. Let  $Y = Y_1 + Y_2$ . Then has range  $\mathcal{Y} \equiv [0, 2]$ , and

$$F_Y(y) = P(Y \leq y) = P(Y_1 + Y_2 \leq y)$$

To compute  $P(Y_1 + Y_2 \leq y)$ , we need to integrate  $f_{Y_1, Y_2}$  over the set

$$A_y = \{(y_1, y_2) : 0 < y_1, y_2 < 1, y_1 + y_2 \leq y\}$$

This region is a portion of the unit square (that is,  $\mathcal{Y}_1 \times \mathcal{Y}_2$ ); the line  $y_1 + y_2 = y$  is a line with negative slope that cuts the  $y_1$  (horizontal) axis at  $y_1 = y$ , and the  $y_2$  axis (vertical) at  $y_2 = y$ . The region is plotted in the figure below: the dotted line is line  $y_1 + y_2 = y$ , for  $0 < y < 1$  (left panel) and for  $1 \leq y < 2$  (right panel).



- For  $0 \leq y \leq 1$ ,  $A_y$  is the grey shaded lower triangle in the left panel of the figure. Hence

$$P(Y_1 + Y_2 < y) = \int_0^y \int_0^{y-y_2} 1 \, dy_1 dy_2 = \int_0^y (y - y_2) dy_2 = \frac{y^2}{2}.$$

- For  $1 \leq y \leq 2$ ,  $A_y$  is more complicated see the figure (right panel). It is easier mathematically to describe the **complement** of  $A_y$  within  $\mathcal{Y}_1 \times \mathcal{Y}_2$  (the red shaded region right panel of the figure), so we instead compute the complement probability as follows:

$$\begin{aligned} P(Y_1 + Y_2 \leq y) &= 1 - P(Y_1 + Y_2 > y) \\ &= 1 - \int_{y-1}^1 \int_{y-y_2}^1 1 \, dy_1 dy_2 = 1 - \int_{y-1}^1 (1 - y + y_2) dy_2 = -\frac{y^2}{2} + 2y - 1 \end{aligned}$$

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

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

and zero otherwise.

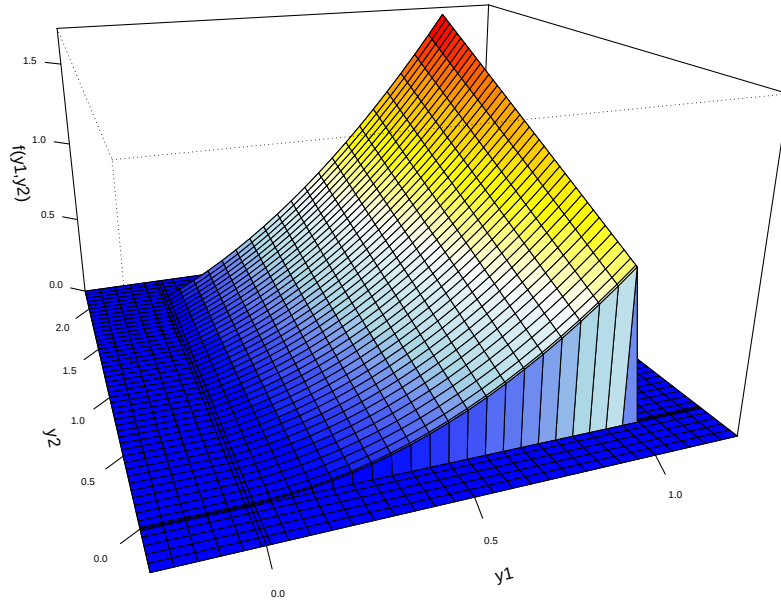
**EXAMPLE 4** Let  $Y_1$  and  $Y_2$  be continuous random variables with ranges  $\mathcal{Y}_1 = (0, 1)$ ,  $\mathcal{Y}_2 = (0, 2)$  and joint pdf

$$f_{Y_1, Y_2}(y_1, y_2) = c \left( y_1^2 + \frac{y_1 y_2}{2} \right) \quad 0 < y_1 < 1, 0 < y_2 < 2$$

and zero otherwise. To calculate  $c$ , we have

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_1 dy_2 &= \int_0^2 \left\{ \int_0^1 c \left( y_1^2 + \frac{y_1 y_2}{2} \right) dy_1 \right\} dy_2 = \int_0^2 c \left[ \frac{y_1^3}{3} + \frac{y_1^2 y_2}{4} \right]_0^1 dy_2 \\ &= \int_0^2 c \left( \frac{1}{3} + \frac{y_2}{4} \right) dy_2 \\ &= c \left[ \frac{y_2}{3} + \frac{y_2^2}{8} \right]_0^2 = c \frac{7}{6} \end{aligned}$$

so  $c = 6/7$ .



The marginal pdf of  $Y_1$  is given, for  $0 < y_1 < 1$ , by

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_2 = \int_0^2 \frac{6}{7} \left( y_1^2 + \frac{y_1 y_2}{2} \right) dy_2 = \frac{6}{7} \left[ y_1^2 y_2 + \frac{y_1 y_2^2}{4} \right]_0^2 = \frac{6y_1(2y_1 + 1)}{7}$$

and is zero otherwise.

To compute  $P(Y_1 > Y_2)$ , let  $A = \{ (y_1, y_2) : 0 < y_1 < 1, 0 < y_2 < 2, y_2 < y_1 \}$ , so that

$$\begin{aligned} P(Y_1 > Y_2) &= \iint_A f_{Y_1, Y_2}(y_1, y_2) dy_2 dy_1 = \int_0^1 \left\{ \int_0^{y_1} \frac{6}{7} \left( y_1^2 + \frac{y_1 y_2}{2} \right) dy_2 \right\} dy_1 \\ &= \int_0^1 \left[ y_1^2 y_2 + \frac{y_1 y_2^2}{4} \right]_0^{y_1} dy_1 \\ &= \int_0^1 \left( y_1^3 + \frac{y_1^3}{4} \right) dy_1 \\ &= \frac{6}{7} \left[ \frac{5y_1^4}{16} \right]_0^1 = \frac{15}{56} \end{aligned}$$

**EXAMPLE 5** Let  $Y_1$ ,  $Y_2$  and  $Y_3$  be continuous random variables with joint ranges

$$\mathcal{Y}^{(3)} = \{(y_1, y_2, y_3) : 0 < y_1 < y_2 < y_3 < 1\}$$

and joint pdf defined by

$$f_{Y_1, Y_2, Y_3}(y_1, y_2, y_3) = c \quad 0 < y_1 < y_2 < y_3 < 1$$

and zero otherwise. To calculate  $c$ , integrate over  $\mathcal{Y}^{(3)}$  as follows:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3}(y_1, y_2, y_3) dy_1 dy_2 dy_3 = 1$$

gives that

$$\int_0^1 \left\{ \int_0^{y_3} \left\{ \int_0^{y_2} c dy_1 \right\} dy_2 \right\} dy_3 = 1$$

Now

$$\int_0^1 \left\{ \int_0^{y_3} \left\{ \int_0^{y_2} c dy_1 \right\} dy_2 \right\} dy_3 = \int_0^1 \left\{ \int_0^{y_3} c y_2 dy_2 \right\} dy_3 = \int_0^1 \frac{c y_3^2}{2} dy_3 = \frac{c}{6}$$

and hence  $c = 6$ .

Also, for  $0 < y_3 < 1$ ,  $f_{Y_3}$  is given by

$$f_{Y_3}(y_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3}(y_1, y_2, y_3) dy_1 dy_2 = \int_0^{y_3} \left\{ \int_0^{y_2} 6 dy_1 \right\} dy_2 = \int_0^{y_3} 6 y_2 dy_2 = 3 y_3^2$$

and is zero otherwise. Similar calculations for  $Y_1$  and  $Y_2$  give

$$f_{Y_1}(y_1) = 3(1 - y_1)^2 \quad 0 < y_1 < 1$$

$$f_{Y_2}(y_2) = 6y_2(1 - y_2) \quad 0 < y_2 < 1$$

with both densities equal to zero outside of these ranges.

Furthermore, for the joint pdf of  $Y_1$  and  $Y_2$ , we have

$$f_{Y_1, Y_2}(y_1, y_2) = \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3}(y_1, y_2, y_3) dy_3 = \int_{y_2}^1 6 dy_3 = 6(1 - y_2) \quad 0 < y_1 < y_2 < 1$$

and zero otherwise. Combining these results, we have, for example, for the conditional pdf of  $Y_1$  given  $Y_2 = y_2$ ,

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1, Y_2}(y_1, y_2)}{f_{Y_2}(y_2)} = \frac{1}{y_2} \quad 0 < y_1 < y_2$$

and zero otherwise for **fixed**  $y_2$ .

**EXAMPLE 6** Let  $Y_1, Y_2$  be continuous random variables with joint pdf given by

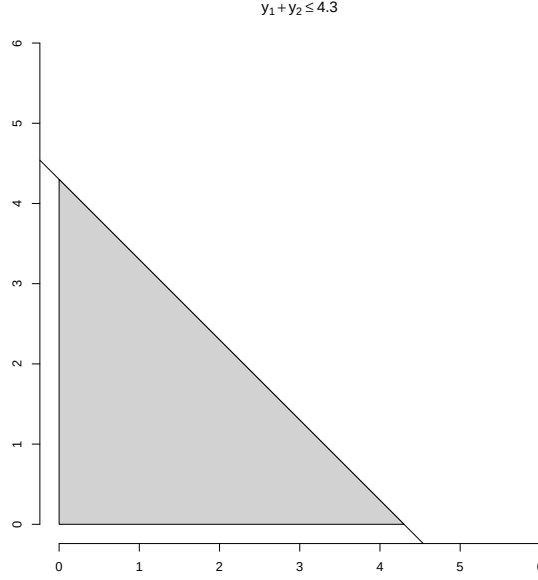
$$f_{Y_1, Y_2}(y_1, y_2) = y_1 \exp\{-(y_1 + y_2)\} \quad y_1, y_2 > 0$$

and zero otherwise. Thus, for  $y_1, y_2 > 0$ ,

$$f_{Y_1, Y_2}(y_1, y_2) = y_1 \exp\{-y_1\} \times \exp\{-y_2\} = f_{Y_1}(y_1)f_{Y_2}(y_2)$$

and  $Y_1$  and  $Y_2$  are independent.

Let  $Y = Y_1 + Y_2$ . For  $y > 0$  let  $A_y = \{(y_1, y_2) : y_1 + y_2 \leq y\}$ . Then the region  $A_y$  is the region in grey in the figure, for the value  $y = 4.3$ :



Then

$$F_Y(y) = P(Y \leq y) = P(Y_1 + Y_2 \leq y) = \iint_{A_y} f_{Y_1, Y_2}(y_1, y_2) dy_1 dy_2 = \int_0^y \left\{ \int_0^{y-y_1} f_{Y_1, Y_2}(y_1, y_2) dy_2 \right\} dy_1$$

We substitute  $t = y_2 + y_1$  (so that  $y_2 = t - y_1$ ) in the inner integral to obtain

$$F_Y(y) = \int_0^y \left\{ \int_0^y f_{Y_1, Y_2}(y_1, t - y_1) dt \right\} dy_1$$

Differentiating under the outer integral sign

$$\begin{aligned} f_Y(y) &= \frac{F_Y(y)}{dy} = \frac{d}{dy} \left\{ \int_0^y \left\{ \int_0^y f_{Y_1, Y_2}(y_1, t - y_1) dt \right\} dy_1 \right\} \\ &= \int_0^y \left\{ \frac{d}{dy} \left\{ \int_0^y f_{Y_1, Y_2}(y_1, t - y_1) dt \right\} \right\} dy_1 = \int_0^y f_{Y_1, Y_2}(y_1, y - y_1) dy_1 \end{aligned}$$

where the last step follows by the fundamental rule of calculus. Thus

$$f_Y(y) = \int_0^y f_{Y_1, Y_2}(y_1, y - y_1) dy_1 = \int_0^y y_1 \exp\{-(y_1 + (y - y_1))\} dy_1 = \frac{1}{2}y^2 e^{-y} \quad y > 0$$

and zero otherwise. Thus, in fact,  $Y \sim \text{Gamma}(3, 1)$ ; in this case, we can see that  $Y_1$  and  $Y_2$  are independent, with  $Y_1 \sim \text{Gamma}(2, 1)$  and  $Y_2 \sim \text{Gamma}(1, 1)$ .

**EXAMPLE 7** Let  $Y_1, Y_2$  be continuous random variables with joint pdf given by

$$f_{Y_1, Y_2}(y_1, y_2) = 2(y_1 + y_2) \quad 0 \leq y_1 \leq y_2 \leq 1$$

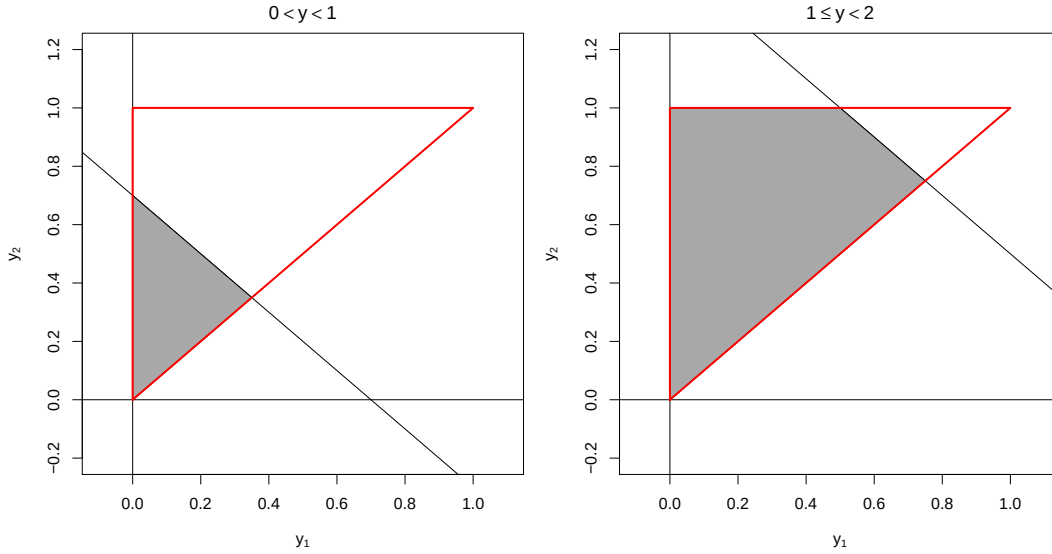
and zero otherwise. Let  $Y = Y_1 + Y_2$ ; clearly  $Y$  takes values on  $\mathcal{Y} \equiv [0, 2]$ ; If  $Y_1$  can take values on the real line, then

$$F_Y(y) = P(Y \leq y) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{y-y_1} f_{Y_1, Y_2}(y_1, y_2) dy_2 \right\} dy_1$$

and changing variables, and then differentiating under the integral sign as before yields

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y - y_1) dy_1,$$

However, the joint pdf is positive only on a subset of  $\mathbb{R}^2$ , that is, the region outlined in red in the figure below: the grey shaded region is the portion of the support of the joint pdf for which  $y_1 + y_2 \leq y$ , with  $y = 0.7$  (left panel) and  $y = 1.5$  (right panel).



We have that  $f_{Y_1, Y_2}(y_1, y - y_1) = 2y$ ; this holds when both  $y_1$  and  $y - y_1$  lie in the interval  $[0, 1]$  with  $y_1 \leq y - y_1$  for fixed  $y$ .

- For  $0 \leq y \leq 1$ , the constraints

$$0 \leq y_1 \leq y - y_1 \leq 1$$

imply that  $0 \leq 2y_1 \leq y$ , that is,  $0 \leq y_1 \leq y/2$ ;

- For  $1 \leq y \leq 2$  the constraints imply  $1 - y \leq y_1 \leq y/2$ .

Hence

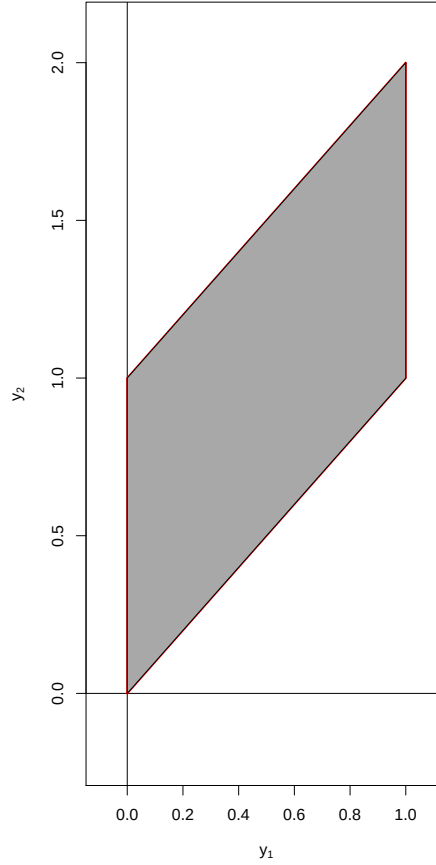
$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y - y_1) dy_1 = \begin{cases} \int_0^{y/2} 2y dy_1 = y^2 & 0 \leq y \leq 1 \\ \int_{y-1}^{y/2} 2y dy_1 = y(2 - y) & 1 \leq y \leq 2 \end{cases}$$

and zero otherwise. It is straightforward to check that this density is a valid pdf.

**EXAMPLE 8** Let  $Y_1, Y_2$  be continuous random variables with joint pdf given by

$$f_{Y_1, Y_2}(y_1, y_2) = c \quad 0 < y_1 < 1, y_1 < y_2 < y_1 + 1$$

and zero otherwise. The support of the joint pdf is the grey shaded region.



To calculate  $c$ , we have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_2 dy_1 = \int_0^1 \int_{y_1}^{y_1+1} c dy_2 dy_1 = \int_0^1 c [y_2]_{y_1}^{y_1+1} dy_1 = \int_0^1 c dy_1 = c$$

so  $c = 1$ . The marginal pdf of  $Y_1$  is given by

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_2 = \int_{y_1}^{y_1+1} 1 dy_2 = 1 \quad 0 < y_1 < 1$$

and zero otherwise, and the marginal pdf for  $Y_2$  is given by

$$f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_1 = \begin{cases} \int_0^{y_2} 1 dy_1 & = y_2 & 0 < y_2 < 1 \\ \int_{y_2-1}^1 1 dy_1 & = 2 - y_2 & 1 \leq y_2 < 2 \end{cases}$$

and zero otherwise. Note that the support in the figure confirms that the range of integration when we marginalize over  $y_1$  (in the horizontal strip at a fixed  $y_2$  value) depends on the value of  $y_2$  chosen.