$$= \int_{t_1=-\infty}^{x_1} \int_{t_2=-\infty}^{x_2} \cdots \int_{t_p=-\infty}^{x_p} f(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1, \forall x = (x_1, x_2, \cdots, x_p) \in \mathbb{R}^p.$$

The function f is called the joint probability density function (joint p.d.f.) of X.

Remark 1.305. Let X be a continuous random vector with joint DF F_X and joint p.d.f. f_X. Then we have the following observations.

(a) F_X is jointly continuous in all co-ordinates.

(b) $\mathbb{P}(X = x) = 0, \forall x \in \mathbb{R}^p$. More generally, if $A \subset \mathbb{R}^p$ is finite or countably infinite, then by the finite/countable additivity of $\mathbb{P}_X$, we have

$$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = \sum_{x \in A} \mathbb{P}_X(\{x\}) = \sum_{x \in A} \mathbb{P}(X = x) = 0.$$

(c) By definition, we have $f_X(x) \geq 0, \forall x \in \mathbb{R}^p$ and

$$\begin{aligned} 1 &= \lim_{x_j \to \infty \, \forall j} F_X(x_1, x_2, \cdots, x_p) \\ &= \lim_{x_j \to \infty \, \forall j} \int_{t_1=-\infty}^{x_1} \int_{t_2=-\infty}^{x_2} \cdots \int_{t_p=-\infty}^{x_p} f(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1. \end{aligned}$$

(d) Suppose that the joint p.d.f. f_X of a p-dimensional random vector X is piecewise continuous. Then by the Fundamental Theorem of Calculus (specifically multivariable Calculus), we have

$$f_X(x_1, x_2, \cdots, x_p) = \frac{\partial^p}{\partial x_1 \partial x_2 \cdots \partial x_p} F_X(x_1, x_2, \cdots, x_p),$$

wherever the partial derivative on the right hand side exists.

(e) If X is a p-dimensional random vector such that its joint DF F_X is continuous on $\mathbb{R}^p$ and such that the partial derivative $\frac{\partial^p}{\partial x_1 \partial x_2 \cdots \partial x_p} F_X$ exists everywhere except possibly on a countable number of curves on $\mathbb{R}^p$. Let $A \subset \mathbb{R}^p$ denote the set of all points on such curves. Then X is a continuous random vector with the joint p.d.f.

$$f_X(x) = \begin{cases} \frac{\partial^p}{\partial x_1 \partial x_2 \cdots \partial x_p} F_X(x), \text{ if } x = (x_1, x_2, \cdots, x_p) \in A^c, \\ 0, \text{ if } x = (x_1, x_2, \cdots, x_p) \in A. \end{cases}$$

(f) The joint p.d.f. of a continuous random vector is not unique. As in the case of continuous RVs, the joint p.d.f. is determined uniquely upto sets of 'volume 0'. Here, we also get versions of the joint p.d.f..

(g) For $A \subset \mathbb{R}^p$, we have

$$\begin{aligned}\mathbb{P}(X \in A) &= \iiint_A f_X(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1 \\ &= \iiint_{\mathbb{R}^p} f_X(t_1, t_2, \cdots, t_p) 1_A(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1,\end{aligned}$$

provided the integral can be defined. We do not prove this statement in this course.

(h) For any $j \in \{1, 2, \cdots, p\}$, for $x_j \in \mathbb{R}$

$$\begin{aligned}F_{X_j}(x_j) &= \mathbb{P}(X_j \in (-\infty, x_j]) \\ &= \mathbb{P}(X_1 \in \mathbb{R}, \cdots, X_{j-1} \in \mathbb{R}, X_j \in (-\infty, x_j], X_{j+1} \in \mathbb{R}, \cdots, X_p \in \mathbb{R}) \\ &= \mathbb{P}(X \in \mathbb{R} \times \cdots \times \mathbb{R} \times (-\infty, x_j] \times \mathbb{R} \times \cdots \times \mathbb{R}) \\ &= \int_{t_1=-\infty}^{\infty} \cdots \int_{t_{j-1}=-\infty}^{\infty} \int_{t_j=-\infty}^{x_j} \int_{t_{j+1}=-\infty}^{\infty} \cdots \int_{t_p=-\infty}^{\infty} f_X(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1.\end{aligned}$$

Consider $g_j : \mathbb{R} \to \mathbb{R}$ defined by

$$g_j(t_j) := \int_{t_1=-\infty}^{\infty} \cdots \int_{t_{j-1}=-\infty}^{\infty} \int_{t_{j+1}=-\infty}^{\infty} \cdots \int_{t_p=-\infty}^{\infty} f_X(t_1, t_2, \cdots, t_p)\, dt_p \cdots dt_{j-1} dt_{j+1} \cdots dt_2 dt_1.$$

It is immediate that g_j satisfies the properties of a p.d.f. and $F_{X_j}(x_j) = \int_{t_j=-\infty}^{x_j} g_j(t_j)\, dt_j$. Therefore, X_j is a continuous RV with p.d.f. g_j. More generally, all marginal distributions of X are also continuous and can be obtained by integrating out the unnecessary co-ordinates. The function g_j is usually referred to as the marginal p.d.f. of X_j.

Remark 1.306. Let $f : \mathbb{R}^p \to [0, \infty)$ be an integrable function with

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1 = 1.$$

Then f is the joint p.d.f. of some p-dimensional continuous random vector X. We are not going to discuss the proof of this statement in this course.

We can identify the independence of the component RVs for a continuous random vector via the joint p.d.f.. The proof is similar to Theorem 1.294 and is skipped for brevity.

Theorem 1.307. *Let $X = (X_1, X_2, \cdots, X_p)$ be a continuous random vector with joint DF F_X, joint p.d.f. f_X. Let f_{X_j} denote the marginal p.d.f. of X_j. Then $X_1, X_2, \cdots, X_p$ are independent if and only if*

$$f_{X_1,X_2,\cdots,X_p}(x_1, x_2, \cdots, x_p) = \prod_{j=1}^{p} f_{X_j}(x_j), \forall x_1, x_2, \cdots, x_p \in \mathbb{R}.$$

Example 1.308. Given p.d.f.s $f_1, f_2, \cdots, f_p : \mathbb{R} \to [0, \infty)$, consider the function $f : \mathbb{R}^p \to [0, \infty)$ defined by

$$f(x) := \prod_{j=1}^{p} f_j(x_j), \forall x = (x_1, x_2, \cdots, x_p) \in \mathbb{R}^p.$$

Then

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(t_1, t_2, \cdots, t_p)\, dt_p dt_{p-1} \cdots dt_2 dt_1 = 1.$$

By Remark 1.306, we have that f is the joint p.d.f. of a p-dimensional continuous random vector such that the component RVs are independent, by Theorem 1.307. Using this method, we can construct many examples of continuous random vectors.

Remark 1.309. Let $X = (X_1, X_2, \cdots, X_p)$ be a continuous random vector with joint p.d.f. f_X. Then $X_1, X_2, \cdots, X_p$ are independent if and only if

$$f_{X_1,X_2,\cdots,X_p}(x_1, x_2, \cdots, x_p) = \prod_{j=1}^{p} g_j(x_j), \forall x_1, x_2, \cdots, x_p \in \mathbb{R}$$

for some integrable functions $g_1, g_2, \cdots, g_p : \mathbb{R} \to [0, \infty)$. In this case, the marginal p.d.fs f_{X_j} have the form $c_j g_j$, where the number c_j can be determined from the relation $c_j = \left(\int_{-\infty}^{\infty} g_j(x)\, dx\right)^{-1}$.

Example 1.310. Let $Z = (X, Y)$ be a 2-dimensional continuous random vector with the joint p.d.f. of the form

$$f_Z(x, y) = \begin{cases} \alpha xy, & \text{if } 0 < x < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

for some constant $\alpha \in \mathbb{R}$. For f_Z to take non-negative values, we must have $\alpha > 0$. Now,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_Z(x,y)\,dxdy = \int_{y=0}^{1}\int_{x=0}^{y} \alpha xy\,dxdy = \int_{y=0}^{1} \alpha\frac{y^3}{2}\,dy = \frac{\alpha}{8}.$$

For f_Z to be a joint p.d.f., we need $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_Z(x,y)\,dxdy = 1$ and hence $\alpha = 8 > 0$. Also note that for this value of α, f_Z takes non-negative values. The marginal p.d.f. f_X of X can now be computed as follows.

$$f_X(x) = \int_{-\infty}^{\infty} f_Z(x,y)\,dy = \begin{cases} \int_{y=x}^{1} 8xy\,dy, \text{ if } x \in (0,1) \\ 0, \text{ otherwise} \end{cases} = \begin{cases} 4x[1-x^2], \text{ if } x \in (0,1) \\ 0, \text{ otherwise} \end{cases}.$$

The marginal p.d.f. f_Y of Y follows by a similar computation.

$$f_Y(y) = \int_{-\infty}^{\infty} f_Z(x,y)\,dx = \begin{cases} \int_{x=0}^{y} 8xy\,dx, \text{ if } y \in (0,1) \\ 0, \text{ otherwise} \end{cases} = \begin{cases} 4y^3, \text{ if } y \in (0,1) \\ 0, \text{ otherwise} \end{cases}.$$

Observe that $f_Z(\frac{1}{2},\frac{1}{2}) = 0$ and $f_X(\frac{1}{2})f_Y(\frac{1}{2}) = \frac{3}{2} \times \frac{1}{2} = \frac{3}{4}$. Hence X and Y are not independent.

Example 1.311. Let $U = (X,Y,Z)$ be a 3-dimensional continuous random vector with the joint p.d.f. of the form

$$f_U(x,y,z) = \begin{cases} \alpha xyz, \text{ if } x,y,z \in (0,1) \\ 0, \text{ otherwise} \end{cases}$$

for some constant $\alpha \in \mathbb{R}$. For f_Z to take non-negative values, we must have $\alpha > 0$. Now,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_U(x,y,z)\,dxdydz = \int_{x=0}^{1}\int_{y=0}^{1}\int_{z=0}^{1} \alpha xyz\,dxdydz = \frac{\alpha}{8}.$$

For f_U to be a joint p.d.f., we need $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_U(x,y,z)\,dxdydz = 1$ and hence $\alpha = 8 > 0$. Also note that for this value of α, f_U takes non-negative values. The marginal p.d.f. f_X of X can now be computed as follows.

$$f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_U(x,y,z)\,dydz = \begin{cases} \int_{z=0}^{1}\int_{y=0}^{1} 8xyz\,dydz, \text{ if } x \in (0,1) \\ 0, \text{ otherwise} \end{cases} = \begin{cases} 2x, \text{ if } x \in (0,1) \\ 0, \text{ otherwise} \end{cases}.$$

By the symmetry of $f_U(x,y,z)$ in the variables x, y and z, we conclude that $X \stackrel{d}{=} Y \stackrel{d}{=} Z$. Observe that $f_{X,Y,Z}(x,y,z) = f_X(x)f_Y(y)f_Z(z), \forall x,y,z$ and hence the RVs X, Y, Z are independent.

Note 1.312. There are random vectors which are neither discrete nor continuous. We do not discuss such examples in this course.

Remark 1.313 (Conditional Distribution for continuous random vectors). We now discuss an analogue of conditional distributions as discussed in Remark 1.299 for discrete random vectors. To avoid notational complexity, we work in dimension 2. Let (X,Y) be a 2-dimensional continuous random vector with joint DF $F_{X,Y}$ and joint p.d.f. $f_{X,Y}$. Let f_X and f_Y denote the marginal p.d.fs of X and Y respectively. Since $\mathbb{P}(X = x) = 0, \forall x \in \mathbb{R}$, expressions of the form $\mathbb{P}(Y \in A \mid X = x)$ are not defined for $A \subset \mathbb{R}$. We consider $x \in \mathbb{R}$ such that $f_X(x) > 0$ and look at the following computation. For $y \in \mathbb{R}$,

$$\begin{aligned}
\lim_{h\downarrow 0} \mathbb{P}(Y \leq y \mid x - h < X \leq x) &= \lim_{h\downarrow 0} \frac{\mathbb{P}(Y \leq y, x - h < X \leq x)}{\mathbb{P}(x - h < X \leq x)} \\
&= \lim_{h\downarrow 0} \frac{\int_{x-h}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\, dsdt}{\int_{x-h}^{x} f_X(t)\, dt} \\
&= \lim_{h\downarrow 0} \frac{\frac{1}{h}\int_{x-h}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\, dsdt}{\frac{1}{h}\int_{x-h}^{x} f_X(t)\, dt} \\
&= \frac{\int_{-\infty}^{y} f_{X,Y}(x,s)\, ds}{f_X(x)} \\
&= \int_{-\infty}^{y} \frac{f_{X,Y}(x,s)}{f_X(x)}\, ds
\end{aligned}$$

Here, we have assumed continuity of the p.d.fs. Motivated by the above computation, we define the conditional DF of Y given $X = x$ (provided $f_X(x) > 0$) by

$$F_{Y|X}(y \mid x) := \lim_{h\downarrow 0} \mathbb{P}(Y \leq y \mid x - h < X \leq x),\ y \in \mathbb{R}$$

and the conditional p.d.f. of Y given $X = x$ (provided $f_X(x) > 0$) by

$$f_{Y|X}(y \mid x) := \frac{f_{X,Y}(x,y)}{f_X(x)},\ y \in \mathbb{R}.$$

These calculations generalizes to the higher dimensions as follows. Let $X = (X_1, X_2, \cdots, X_{p+q})$ be a continuous random vector with joint p.d.f. f_X. Let $Y = (X_1, X_2, \cdots, X_p)$ and $Z = (X_{p+1}, X_{p+2}, \cdots, X_{p+q})$. If $z \in \mathbb{R}^q$ be such that $f_Z(z) > 0$, then we define the conditional DF of Y given $Z = z$ by

$$F_{Y|Z}(y \mid z) := \lim_{\substack{h_j \downarrow 0 \\ j=p+1,p+2,\cdots,p+q}} \mathbb{P}(X_1 \le y_1, \cdots, X_p \le y_p \mid x_j - h_j < X_j \le x_j, \forall j),\ y \in \mathbb{R}^p$$

and the conditional p.d.f. of Y given $Z = z$ by

$$f_{Y|Z}(y \mid z) := \frac{f_{Y,Z}(y,z)}{f_Z(z)},\ y \in \mathbb{R}^p.$$

Note 1.314. For notational convenience, we have discussed the conditional distribution of first p component RVs with respect to the final q component RVs. However, as long as the $(p+q)$-dimensional joint distribution is known, we can discuss the conditional distribution of any of the k-component RVs with respect to the other $(p+q-k)$-component RVs.

Note 1.315. Let (X, Y) be a 2-dimensional continuous random vector such that X and Y are independent. Then $f_{X,Y}(x,y) = f_X(x) f_Y(y), \forall x, y \in \mathbb{R}$. Then

$$f_{Y|X}(y \mid x) = f_Y(y), \forall y \in \mathbb{R},$$

provided $f_X(x) > 0$. This statement can be generalized to higher dimensions with appropriate changes in the notation.

Example 1.316. In Example 1.310, we have, for fixed $x \in (0,1)$,

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{2xy}{x(1-x^2)}, & \text{if } y \in (x,1) \\ 0, & \text{otherwise.} \end{cases}$$

Earlier in Week 5, we have discussed about the distribution of functions of RVs. We now generalize the same concept for random vectors.

Remark 1.317. Let $X = (X_1, \ldots, X_p)$ be a p-dimensional discrete/continuous random vector with joint p.m.f./p.d.f. f_X. We are interested in the distribution of $Y = h(X)$ for functions $h : \mathbb{R}^p \to$

$\mathbb{R}^q$. Here, $Y = (Y_1, \ldots, Y_q)$ is a q-dimensional random vector with $Y_j = h_j(X_1, \ldots, X_p)$, where $h_j : \mathbb{R}^p \to \mathbb{R}, j = 1, 2, \cdots, q$ denotes the component functions of h. The distribution of Y is uniquely determined as soon as we are able to compute the joint DF F_Y of Y. Note that

$$F_Y(y_1, \cdots, y_q) = \mathbb{P}(Y_1 \le y_1, \cdots, Y_q \le y_q) = \mathbb{P}(h_1(X) \le y_1, \cdots, h_q(X) \le y_q), \forall (y_1, \cdots, y_q) \in \mathbb{R}^q.$$

Once the joint DF F_Y is known, the joint p.m.f./p.d.f. of Y can then be deduced by standard techniques.

Example 1.318. Let $X_1 \sim Uniform(0,1)$ and $X_2 \sim Uniform(0,1)$ be independent RVs. Suppose we are interested in the distribution of $Y = X_1 + X_2$. By independence of X_1 and X_2, the joint p.d.f. (X_1, X_2) is given by

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2}(x_2) \\ &= \begin{cases} 1, & \text{if } x_1, x_2 \in (0,1) \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Consider the function $h : \mathbb{R}^2 \to \mathbb{R}$ defined by $h(x_1, x_2) := x_1 + x_2, \forall (x_1, x_2) \in \mathbb{R}^2$. Then $Y = h(X_1, X_2)$. Now, for $y \in \mathbb{R}$

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) \\ &= \mathbb{P}(h(X_1, X_2) \le y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_{(-\infty, y]}(h(x_1, x_2)) f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2 \\ &= \int_0^1 \int_0^1 1_{(-\infty, y]}(x_1 + x_2)\, dx_1 dx_2 \\ &= \begin{cases} 0, \text{ if } y < 0, \\ \int_{x_1=0}^{y} \int_{x_2=0}^{y-x_1} dx_2 dx_1, \text{ if } 0 \le y < 1, \\ 1 - \frac{1}{2} \times (2-y) \times (2-y), \text{ if } 1 \le y < 2, \\ 1, \text{ if } y \ge 2 \end{cases} \end{aligned}$$

$$= \begin{cases} 0, \text{ if } y < 0, \\ \frac{y^2}{2}, \text{ if } 0 \le y < 1, \\ \frac{4y-y^2-2}{2}, \text{ if } 1 \le y < 2, \\ 1, \text{ if } y \ge 2 \end{cases}$$

Here, F_Y is differentiable everywhere except possibly at the points $0, 1, 2$ and

$$F'_Y(y) = \begin{cases} y, \text{ if } y \in (0,1), \\ 2-y, \text{ if } y \in (1,2), \\ 0, \text{ otherwise.} \end{cases}$$

Observe that $\int_{-\infty}^{\infty} F'_Y(y)\,dy = 1$ and the derivative is non-negative. Hence, Y is a continuous RV with the p.d.f. given by F'_Y.

As done in the case of RVs, in the setting of Remark 1.317, we consider the computation of the joint p.m.f./p.d.f. of Y directly, instead of computing the joint DF F_Y first. The next result is a direct generalization of Theorem 1.164 and we skip the proof for brevity.

Theorem 1.319 (Change of Variables for Discrete random vectors)**.** *Let $X = (X_1, \ldots, X_p)$ be a p-dimensional discrete random vector with joint p.m.f. f_X and support S_X. Let $h = (h_1, \cdots, h_q) : \mathbb{R}^p \to \mathbb{R}^q$ be a function and let $Y = (Y_1, \cdots, Y_q) = h(X) = (h_1(X), \cdots, h_q(X))$. Then Y is a discrete random vector with support*

$$S_Y = h(S_X) = \{h(x) : x \in S_X\},$$

joint p.m.f.

$$f_Y(y) = \begin{cases} \displaystyle\sum_{\substack{x \in S_X \\ h(x)=y}} f_X(x), \; \textit{if } y \in S_Y, \\ 0, \; \textit{otherwise} \end{cases}$$

and joint DF

$$F_Y(y) = \sum_{\substack{x \in S_X \\ h(x) \leq y}} f_X(x), \forall y \in \mathbb{R}^q.$$

Example 1.320. Fix $p \in (0,1)$ and let $n_1, \cdots, n_q$ be positive integers. Let $X_1, \cdots, X_q$ be independent RVs with $X_i \sim Binomial(n_i, p), i = 1, \cdots, q$. Here, the with p.m.f.s are given by

$$f_{X_i}(x_i) = \begin{cases} \binom{n_i}{x} p^x (1-p)^{n_i - x}, \forall x \in \{0, 1, \cdots, n_i\}, \\ 0, \text{ otherwise} \end{cases}$$

for $i = 1, \cdots, q$. Using independence, the joint p.m.f. is given by

$$f_X(x_1, \cdots, x_q) = \begin{cases} \prod_{i=1}^q \binom{n_i}{x_i} p^{\sum_{i=1}^q x_i} (1-p)^{n - \sum_{i=1}^q x_i}, \forall (x_1, \cdots, x_q) \in \prod_{i=1}^q \{0, 1, \cdots, n_i\}, \\ 0, \text{ otherwise} \end{cases}$$

where $n = n_1 + \cdots + n_q$. Consider $Y = X_1 + \cdots + X_q$. Now, if $y \notin \{0, 1, \cdots, n\}$, $f_Y(y) = \mathbb{P}(X_1 + \cdots + X_q = y) = 0$ and if $y \in \{0, 1, \cdots, n\}$, then

$$\begin{aligned} f_Y(y) &= \mathbb{P}(X_1 + \cdots + X_q = y) \\ &= \sum_{\substack{(x_1, \cdots, x_q) \in \prod_{i=1}^q \{0, 1, \cdots, n_i\} \\ x_1 + \cdots + x_q = y}} f_X(x_1, \cdots, x_q) \\ &= p^y (1-p)^{n-y} \sum_{\substack{(x_1, \cdots, x_q) \in \prod_{i=1}^q \{0, 1, \cdots, n_i\} \\ x_1 + \cdots + x_q = y}} \prod_{i=1}^q \binom{n_i}{x_i} \\ &= \binom{n}{y} p^y (1-p)^{n-y}. \end{aligned}$$

Therefore, $Y = X_1 + \cdots + X_q \sim Binomial(n, p)$ with $n = n_1 + \cdots + n_q$.

Remark 1.321. We had earlier mentioned that $Bernoulli(p)$ distribution is the same as $Binomial(1, p)$ distribution. Using the above computation, we can identify a $Binomial(n, p)$ RV as a sum of n independent RVs each having distribution $Bernoulli(p)$. We shall come back to this observation in later lectures.

For continuous random vectors, we have the following generalization of Theorem 1.172. Proof of this result is being skipped.

Theorem 1.322. *Let $X = (X_1, \ldots, X_p)$ be a p-dimensional continuous random vector with joint p.d.f. f_X. Suppose that $\{x \in \mathbb{R}^p : f_X(x) > 0\}$ can be written as a disjoint union $\cup_{i=1}^k S_i$ of open sets in $\mathbb{R}^p$.*

Let $h^j : \mathbb{R}^p \to \mathbb{R}, j = 1, \cdots, p$ be functions such that $h = (h^1, \cdots, h^p) : S_i \to \mathbb{R}^p$ is one-to-one with inverse $h_i^{-1} = ((h_i^1)^{-1}, \cdots, (h_i^p)^{-1})$ for each $i = 1, \cdots, k$. Moreover, assume that $(h_i^j)^{-1}, i = 1, 2, \cdots, k; j = 1, \cdots, p$ have continuous partial derivatives and the Jacobian determinant of the transformation

$$J_i := \begin{vmatrix} \frac{\partial (h_i^1)^{-1}}{\partial y_1}(t) & \cdots & \frac{\partial (h_i^1)^{-1}}{\partial y_p}(y) \\ \vdots & \vdots & \vdots \\ \frac{\partial (h_i^p)^{-1}}{\partial y_1}(y) & \cdots & \frac{\partial (h_i^p)^{-1}}{\partial y_p}(y) \end{vmatrix} \neq 0, \forall i = 1, \cdots, k.$$

Then the p-dimensional random vector $Y = (Y_1, \cdots, Y_p) = h(X) = (h^1(X), \cdots, h^p(X))$ is a continuous with joint p.d.f.

$$f_Y(y) = \sum_{i=1}^k f_X((h_i^1)^{-1}(y), \cdots, (h_i^p)^{-1}(y)) \, |J_i| \, 1_{h(S_i)}(y).$$

Example 1.323. Fix $\lambda > 0$. Let $X_1 \sim Exponential(\lambda)$ and $X_2 \sim Exponential(\lambda)$ be independent RVs defined on the same probability space. The joint distribution of (X_1, X_2) is given by the joint p.d.f.

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \begin{cases} \frac{1}{\lambda^2} \exp(-\frac{x_1+x_2}{\lambda}), & \text{if } x_1 > 0, x_2 > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Consider the function

$$h(x_1, x_2) = \begin{cases} (x_1 + x_2, \frac{x_1}{x_1+x_2}), \forall x_1 > 0, x_2 > 0, \\ 0, \text{otherwise.} \end{cases}$$

Here, $\{(x_1, x_2) \in \mathbb{R}^2 : f_{X_1,X_2}(x_1, x_2) > 0\} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$ and $h : \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\} \to \mathbb{R}^2$ is one-to-one with range $(0, \infty) \times (0, 1)$. The inverse function is

$h^{-1}(y_1, y_2) = (y_1y_2, y_1(1-y_2))$ for $(y_1, y_2) \in (0, \infty) \times (0,1)$ with Jacobian determinant given by

$$J(y_1, y_2) = \begin{vmatrix} y_2 & y_1 \\ 1-y_2 & -y_1 \end{vmatrix} = -y_1.$$

Now, $Y = (Y_1, Y_2) = h(X_1, X_2) = (X_1 + X_2, \frac{X_1}{X_1+X_2})$ has the joint p.d.f. given by

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= \begin{cases} f_{X_1,X_2}(y_1y_2, y_1(1-y_2))|J(y_1,y_2)|, & \text{if } y_1 > 0, y_2 \in (0,1) \\ 0, & \text{otherwise.} \end{cases} \\ &= \begin{cases} \frac{1}{\lambda^2} y_1 \exp\left(-\frac{y_1}{\lambda}\right), & \text{if } y_1 > 0, y_2 \in (0,1) \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Now, we compute the marginal distributions. The marginal p.d.f. f_{Y_1} is given by

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1, y_2)\, dy_2 = \begin{cases} \frac{1}{\lambda^2} y_1 \exp\left(-\frac{y_1}{\lambda}\right), & \text{if } y_1 > 0 \\ 0, & \text{otherwise} \end{cases}$$

and the marginal p.d.f. f_{Y_2} is given by

$$f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1, y_2)\, dy_1 = \begin{cases} 1, & \text{if } y_2 \in (0,1) \\ 0, & \text{otherwise} \end{cases}.$$

Therefore $Y_1 = X_1 + X_2 \sim Gamma(2, \lambda)$ and $Y_2 = \frac{X_1}{X_1+X_2} \sim Uniform(0,1)$. Moreover,

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1) f_{Y_2}(y_2), \forall (y_1, y_2) \in \mathbb{R}^2$$

and hence Y_1 and Y_2 are independent.

Remark 1.324. We had earlier mentioned that $Exponential(\lambda)$ distribution is the same as $Gamma(1, \lambda)$ distribution. Using the above computation, we can identify a $Gamma(2, \lambda)$ RV as a sum of two independent RVs each having distribution $Gamma(1, \lambda)$. A more general property in this regard is mentioned in practice problem set 8.