IC105: Probability and Statistics **2021-22-M**

Lecture 16: Classification of Random Vectors

Instructor: Dr. Kuldeep Kumar Kataria *Scribe:*

16.0.1. Independent Random Variables

For an arbitary (countable or uncountable) set Δ, let $\{X_\lambda : \lambda \in \Delta\}$ be a family of random variables.

Definition 16.1. *The random variables $X_\lambda, \lambda \in \Delta$ are said to be mutually independent if for any finite subcollection $\{X_{\lambda_1}, X_{\lambda_2}, \ldots, X_{\lambda_p}\}$ in $\{X_\lambda : \lambda \in \Delta\}$*

$$F_{\lambda_1,\lambda_2,\ldots,\lambda_p}(x_1, x_2, \ldots, x_p) = \prod_{i=1}^{p} F_{\lambda_i}(x_i) \; \forall \; \underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p,$$

where $F_{\lambda_1,\lambda_2,\ldots,\lambda_p}(\cdot)$ denotes the joint d.f. of $(X_{\lambda_1}, X_{\lambda_2}, \ldots, X_{\lambda_p})$ and $F_{\lambda_i}(\cdot),\ \ i = 1, 2, \ldots, p$ denotes the marginal d.f. of X_{λ_i}.

The random variables $X_\lambda, \lambda \in \Delta$ are said to be pairwise independent if for any $\lambda_1, \lambda_2 \in \Delta$ $(\lambda_1 \neq \lambda_2)$

$$F_{\lambda_1,\lambda_2}(x_1, x_2) = F_{\lambda_1}(x_1)F_{\lambda_2}(x_2) \; \forall \; \underline{x} = (x_1, x_2) \in \mathbb{R}^2.$$

Remark 16.2. *(a) Random variables $\{X_\lambda, \lambda \in \Delta\}$ are independent iff those in any finite subset of $\{X_\lambda : \lambda \in \Delta\}$ are independent.*

(b) Let $\Delta_1 \subseteq \Delta_2$. Then r.v.s $\{X_\lambda, \lambda \in \Delta_2\}$ are independent $\implies$ r.v.s $\{X_\lambda, \lambda \in \Delta_1\}$ are independent. In particular, if r.v.s in a collection are indepedent then they are pairwise indepedent. The converse may not be true.

Theorem 16.3. *For any positive integer p (≥ 2) the random variables $X_1, X_2, \ldots, X_p$ are independent iff*

$$F(x_1, x_2, \ldots, x_p) = \prod_{i=1}^{p} F_{X_i}(x_i) \; \forall \; \underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p, \tag{16.1}$$

where $F(\cdot)$ is the joint d.f. of $\underline{X} = (X_1, X_2, \ldots, X_p)$.

Proof. Obviously, if $X_1, X_2, \ldots, X_p$ are independent then (16.1) holds. Conversely suppose that (16.1) holds. Consider a subset of $\{X_1, X_2, \ldots, X_p\}$. For simplicity let this subset be $\{X_1, X_2, \ldots, X_q\}$, for some $2 \leq q \leq p$. Thus for $\underline{x} = (x_1, x_2, \ldots, x_q) \in \mathbb{R}^q$ the joint (marginal) d.f. of $(X_1, X_2, \ldots, X_q)$ is

$$G(x_1, x_2, \ldots, x_q) = \lim_{\substack{x_i \to \infty \\ i=q+1,\ldots,p}} F(x_1, x_2, \ldots, x_q, x_{q+1}, \ldots, x_p) = \lim_{\substack{x_i \to \infty \\ i=q+1,\ldots,p}} \prod_{j=1}^{p} F_{X_j}(x_j) = \prod_{j=1}^{q} F_{X_j}(x_j),$$

$\forall \; \underline{x} = (x_1, x_2, \ldots, x_q) \in \mathbb{R}^q$. Here $F_{X_j}(\cdot)$ is the marginal d.f. of X_j, $j = 1, 2, \ldots, q$. $\square$

16.0.2. Discrete Random Vectors

Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be p-dimensional random vector with d.f. $F(\cdot)$.

Definition 16.4. *(a) The random vector $\underline{X} = (X_1, X_2, \ldots, X_p)$ is said to be a discrete random vector if there exists a countable set S (finite or infinite) such that*

$$\Pr(\underline{X} = \underline{x}) > 0 \; \forall \; \underline{x} \in S, \;\; \textit{and} \;\; \Pr(\underline{X} \in S) = 1.$$

The set S is called support of random vector $\underline{X}$ (or of F).

(b) The joint p.m.f. of $\underline{X}$ having support S is defined by

$$f(\underline{x}) = \begin{cases} \Pr(\underline{X} = \underline{x}), & \textit{if } \underline{x} \in S, \\ 0, & \textit{otherwise}. \end{cases}$$

Remark 16.5. *(a) Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional discrete random vector with p.m.f. $f(\cdot)$ and d.f. $F(\cdot)$ and support S. Then, for any $A \subseteq \mathbb{R}^p$*

$$\Pr(\underline{X} \in A) = \Pr(\underline{X} \in A \cap S) = \sum_{\underline{x} \in A \cap S} f(\underline{x}), \;\; (\Pr(\underline{X} \in S) = 1, \; A \cap S \subseteq S \textit{ and thus } A \cap S \textit{ is a countable set}).$$

Moreover, $F(\underline{x}) = \displaystyle\sum_{\underline{y} \in S \cap (-\infty, \underline{x}]} f(\underline{y}), \; \underline{x} \in \mathbb{R}^p.$

(b) Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional discrete random vector with p.m.f. $f(\cdot)$ and support S. Then the p.m.f. $f : \mathbb{R}^p \to \mathbb{R}$ satisfies:

$$(i) \quad f(\underline{x}) > 0 \; \forall \; x \in S \;\; \textit{and} \; f(\underline{x}) = 0 \; \forall \; x \in S^c, \quad (ii) \; \sum_{\underline{x} \in S} f(\underline{x}) = 1.$$

Conversely suppose that $g : \mathbb{R}^p \to \mathbb{R}$ is a function such that for some countable set T

$$(i) \quad g(\underline{x}) > 0 \; \forall \; x \in T \;\; \textit{and} \; g(\underline{x}) = 0 \; \forall \; x \in T^c, \quad (ii) \; \sum_{\underline{x} \in T} g(\underline{x}) = 1.$$

Then $g(\cdot)$ is a p.m.f. of some p-dimensional discrete random vector having support T.

(c) Marginal distributions of discrete random vector are discrete.

Theorem 16.6. *Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional discrete random vector with p.m.f. $f(\cdot)$ and support S. Then the marginal distribution of any subset of $\{X_1, X_2, \ldots, X_p\}$ (say that of $\underline{Y} = (X_1, X_2, \ldots, X_q), \;\; 1 \leq q < p$) is again discrete with p.m.f.*

$$g(x_1, x_2, \ldots, x_q) = \begin{cases} \sum_{x_{q+1}} \sum_{x_{q+2}} \cdots \sum_{x_p} f(\underline{x}), & \textit{if } \underline{x} \in T, \\ 0, & \textit{otherwise} \end{cases}$$

and support

$$T = \{\underline{y} = (y_1, y_2, \ldots, y_q) \in \mathbb{R}^q : (y_1, y_2, \ldots, y_q, y_{q+1}, \ldots, y_p) \in S, \textit{ for some } (y_{q+1}, y_{q+2}, \ldots, y_p) \in \mathbb{R}^{p-q}\}.$$

Proof. Follows using theorem of total probability. $\square$

Conditional distribution of discrete random vectors

Let $\underline{Y} = (Y_1, Y_2, \ldots, Y_p)$, $\underline{Z} = (Z_1, Z_2, \ldots, Z_q)$ and $\underline{X} = (\underline{Y}, \underline{Z}) = (Y_1, Y_2, \ldots, Y_p, Z_1, Z_2, \ldots, Z_q)$ be random vectors with p.m.f. $f_1,\ f_2$ and f, respectively. Suppose $\underline{X},\ \underline{Y}$ and $\underline{Z}$ have support $S,\ S_1,\ S_2$, respectively. For fixed $\underline{z} \in S_2$ define

$$T_{\underline{z}} = \{\underline{y} = (y_1, y_2, \ldots, y_p) \in \mathbb{R}^p : (\underline{y}, \underline{z}) \in S\}.$$

For fixed $\underline{z} \in S_2$, the conditional p.m.f. of $\underline{Y}$ given $\underline{Z} = \underline{z}$ is defined by

$$f(\underline{y}|\underline{z}) = \Pr(\underline{Y} = \underline{y}|\underline{Z} = \underline{z}) = \frac{\Pr(\underline{X} = (\underline{y}, \underline{z}))}{\Pr(\underline{Z} = \underline{z})} = \begin{cases} \dfrac{f(\underline{y}, \underline{z})}{f_2(\underline{z})},\ \underline{y} \in T_{\underline{z}}, \\ 0,\ \text{otherwise.} \end{cases}$$

Clearly for each $\underline{z} \in S_2$, $f(\cdot|\underline{z})$ is a proper p.m.f. with support $T_{\underline{z}}$. Also for fix $\underline{z} \in S_2$

$$\Pr(Y_1 \le y_1, \ldots, Y_q \le y_q|\underline{Z} = \underline{z}) = \frac{\Pr(Y_1 \le y_1, \ldots, Y_q \le y_q, \underline{Z} = \underline{z})}{\Pr(\underline{Z} = \underline{z})} = \sum_{\underline{s} \in T_{\underline{z}}, \underline{s} \le \underline{y}} \frac{f(\underline{s}, \underline{z})}{f_{\underline{Z}}(\underline{z})} = \sum_{\underline{s} \in T_{\underline{z}}, \underline{s} \le \underline{y}} f(\underline{s}|\underline{z}).$$

Theorem 16.7. *Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional random vector with support S and p.m.f. $f(\cdot)$. Let $f_i(\cdot)$ denote the marginal p.m.f. of X_i, $i = 1, 2, \ldots, p$. Then $X_1, \ldots, X_p$ are independent iff*

$$f(x_1, x_2, \ldots, x_p) = \prod_{i=1}^{p} f_i(x_i)\ \forall\ \underline{x} \in S.$$

Proof. (For $p = 2$)

Suppose that $f(x_1, x_2) = f_1(x_1)f_2(x_2)\ \forall\ \underline{x} = (x_1, x_2) \in S$. Then the d.f. of $\underline{X} = (X_1, X_2)$ is

$$F(x_1, x_2) = \sum\sum_{\substack{(y_1, y_2) \in S \\ y_1 \le x_1, y_2 \le x_2}} f(y_1, y_2) = \sum\sum_{\substack{(y_1, y_2) \in S \\ y_1 \le x_1, y_2 \le x_2}} f_1(y_1)f_2(y_2),\ \ (x_1, x_2) \in \mathbb{R}^2.$$

Let S_1 and S_2 be supports of X_1 and X_2 respectively. Then

$$\begin{aligned} S &= \{(y_1, y_2) \in \mathbb{R}^2 : f(y_1, y_2) > 0\} \\ &= \{(y_1, y_2) \in \mathbb{R}^2 : f_1(y_1)f_2(y_2) > 0\} \\ &= \{(y_1, y_2) \in \mathbb{R}^2 : f_1(y_1) > 0 \ \text{and}\ f_2(y_2) > 0\} \\ &= \{y_1 \in \mathbb{R} : f_1(y_1) > 0\} \times \{y_2 \in \mathbb{R} : f_2(y_2) > 0\} = S_1 \times S_2. \end{aligned}$$

Therefore, for $(x_1, x_2) \in \mathbb{R}^2$

$$F(x_1, x_2) = \sum\sum_{\substack{y_1 \in S_1, y_2 \in S_2 \\ y_1 \le x_1, y_2 \le x_2}} f_1(y_1)f_2(y_2) = \left(\sum_{\substack{y_1 \in S_1 \\ y_1 \le x_1}} f_1(y_1)\right)\left(\sum_{\substack{y_2 \in S_2 \\ y_2 \le x_2}} f_2(y_2)\right) = F_1(x_1)F_2(x_2)$$

where F_1 and F_2 are marginal d.f.s of X_1 and X_2 respectively. This imples X_1 and X_2 are independent. Conversely, suppose that X_1 and X_2 are independent. Then $F(y_1, y_2) = F_1(y_1)F_2(y_2)\ \forall\ (y_1, y_2) \in \mathbb{R}^2$. Then, for $(x_1, x_2) \in \mathbb{R}^2$

$$\begin{aligned}
f(x_1, x_2) &= \Pr(X_1 = x_1, X_2 = x_2)\\
&= \Pr\left(\bigcap_{n=1}^{\infty}\left\{x_1 - \frac{1}{n} < X_1 \le x_1, x_2 - \frac{1}{n} < X_2 \le x_2\right\}\right)\\
&= \lim_{n\to\infty} \Pr\left(x_1 - \frac{1}{n} < X_1 \le x_1, x_2 - \frac{1}{n} < X_2 \le x_2\right)\\
&= \lim_{n\to\infty}\left[F(x_1, x_2) - F\left(x_1 - \frac{1}{n}, x_2\right) - F\left(x_1, x_2 - \frac{1}{n}\right) + F\left(x_1 - \frac{1}{n}, x_2 - \frac{1}{n}\right)\right]\\
&= \lim_{n\to\infty}\left[F_1(x_1)F_2(x_2) - F_1\left(x_1 - \frac{1}{n}\right)F_2(x_2) - F_1(x_1)F_2\left(x_2 - \frac{1}{n}\right) + F_1\left(x_1 - \frac{1}{n}\right)F_2\left(x_2 - \frac{1}{n}\right)\right]\\
&= F_1(x_1)F_2(x_2) - F_1(x_1-)F_2(x_2) - F_1(x_1)F_2(x_2-) + F_1(x_1-)F_2(x_2-)\\
&= (F_1(x_1) - F_1(x_1-))F_2(x_2) - (F_1(x_1) - F_1(x_1-))F_2(x_2-)\\
&= (F_1(x_1) - F_1(x_1-))(F_2(x_2) - F_2(x_2-)) = f_1(x_1)f_2(x_2).
\end{aligned}$$

This completes the proof for $p = 2$ case. Similary, it can be proved for other cases. $\square$

Remark 16.8. *(a) If $\underline{X} = (X_1, X_2, \ldots, X_p)$ is a discrete r.v. with support S and X_i has support S_i, $i = 1, 2, \ldots, p$ then $X_1, X_2, \ldots, X_p$ are independent $\implies S = S_1 \times S_2 \times \cdots \times S_p$.*

(b) Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a discrete random vector with support S and p.m.f. $f(\cdot)$. Then $X_1, \ldots, X_p$ are independent iff

$$f(x_1, x_2, \ldots, x_p) = f_1(x_1)f_2(x_2)\cdots f_p(x_p),\ \ \underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p$$

and $S = A_1 \times A_2 \times \cdots \times A_p$ for some functions $f_1, f_2, \ldots, f_p$ and countable set $A_i = \{x \in \mathbb{R} : f_i(x) > 0\}$, $i = 1, 2, \ldots, p$. In that case the marginal p.m.f. of X_i is $f_i(x) = c_i g_i(x)$, $x \in \mathbb{R}$ for some constant c_i such that $\sum_{x\in A_i} c_i g_i(x) = 1$, $i = 1, 2, \ldots, p$.

(c) If $\underline{X} = (Y, Z)$ is a two-dimensional r.v. then Y and Z are indepedent iff $f(y|z) = f_1(y)\ \forall\ y \in \mathbb{R}$ and $z \in \mathbb{R}$ such that $f_2(z) > 0$, here $f(y|z)$ denotes the conditional p.m.f. of Y given $Z = z$ and $f_1(\cdot)$ denotes the marginal p.d.f. of Y.

(d) One can extend Definition 16.1 to define indepedence of a collection of random vectors. Then analogous of Theorem 16.3, Remark 16.5, Theorem 16.6, Theorem 16.7 and (c) above holds for random vectors.

Example 16.9. *Let $\underline{X} = (X_1, X_2, X_3)$ have the joint p.m.f.*

$$f(x_1, x_2, x_3) = \begin{cases} cx_1x_2x_3,\ x_1 = 1, 2,\ \ x_2 = 1, 2, 3,\ \ x_3 = 1, 3,\\ 0,\ \textit{otherwise}.\end{cases}$$

where c is a real constant.

(a) Find the value of c.

(b) Find the marginal p.m.f.s of X_1, X_2 and X_3.

(c) Are X_1, X_2 and X_3 independent.

(d) Find marginal p.m.f. of (X_1, X_3).

(e) Find conditional p.m.f. of X_1 given $(X_2, X_3) = (2, 1)$.

(f) Are X_1 and X_3 independent.

(g) Compute $\Pr(X_1 = X_2 = X_3)$.

Solution: Here the support of random vector $\underline{X}$ is $S_{\underline{X}} = \{1,2\} \times \{1,2,3\} \times \{1,3\}$.

(a) $\sum_{\underline{x} \in S_{\underline{X}}} f(\underline{x}) = 1 \implies c(1+3+2+6+3+9+2+6+4+12+6+18) = 1 \implies c = \frac{1}{72}$. Clearly $f(\underline{x}) \geq 0$ $\forall\, \underline{x} \in \mathbb{R}^3$.

(b) For $x_1 \neq \{1,2\}$, clearly $f_{X_1}(x_1) = 0$. For $x_1 \in \{1,2\}$

$$f_{X_1}(x_1) = \sum_{(x_2,x_3)\in\{1,2,3\}\times\{1,3\}} \frac{x_1 x_2 x_3}{72} = \frac{x_1}{72}\left(\sum_{x_2=1}^{3} x_2\right)\left(\sum_{x_1=1,3} x_3\right) = \frac{x_1}{3}.$$

Thus,

$$f_{X_1}(x_1) = \begin{cases} \dfrac{x_1}{3}, \ x_1 \in \{1,2\}, \\ 0, \ \text{otherwise.} \end{cases}$$

Similarly,

$$f_{X_2}(x_2) = \begin{cases} \dfrac{x_2}{6}, \ x_2 \in 1,2,3, \\ 0, \ \text{otherwise.} \end{cases} ; \quad f_{X_3}(x_3) = \begin{cases} \dfrac{x_3}{4}, \ x_3 \in 1,3, \\ 0, \ \text{otherwise.} \end{cases}$$

(c) Clearly, $f(x_1,x_2,x_3) = f_1(x_1)f_2(x_2)f_3(x_3)$, $(x_1,x_2,x_3) \in \mathbb{R}^3$ and $S_X = \Delta_1 \times \Delta_2 \times \Delta_3$ where $\Delta_1 = \{1,2\}$, $\Delta_2 = \{1,2,3\}$ and $\Delta_3 = \{1,3\}$.

$$f_1(x_1) = \begin{cases} c_1 x_1, \ x_1 \in \Delta_1, \\ 0, \ \text{otherwise.} \end{cases} ; \ f_2(x_2) = \begin{cases} c_2 x_2, \ x_2 \in \Delta_2, \\ 0, \quad \text{otherwise.} \end{cases} ; \ f_3(x_3) = \begin{cases} c_3 x_3, \ x_3 \in \Delta_3, \\ 0, \ \text{otherwise.} \end{cases}$$

Obviously, $c_1 = \frac{1}{3}$, $c_2 = \frac{1}{6}$ and $c_3 = \frac{1}{4}$. Thus X_1, X_2 and X_3 are independent.

(d) Marginal of (X_1, X_3) is $f_{X_1,X_3}(x_1,x_3) = \sum_{x_2} f_{\underline{X}}(x_1,x_2,x_3) = \dfrac{x_1 x_3}{72} \times 6 = \dfrac{x_1 x_3}{12}$. Thus,

$$f_{X_1,X_3}(x_1,x_3) = \begin{cases} \dfrac{x_1 x_3}{12}, \ (x_1,x_3) \in \{1,2\} \times \{1,3\}, \\ 0, \ \text{otherwise} \end{cases}$$

(e) For $x_1 \in \{1,2\}$,

$$\Pr(X_1 = x_1 | X_2 = 2, X_3 = 1) = \frac{\Pr(X_1 = x_1, X_2 = 2, X_3 = 1)}{\Pr(X_2 = 2, X_3 = 1)} = \frac{x_1 2}{72} \Big/ \frac{1}{12} = \frac{x_1}{3}.$$

Thus

$$f_{X_1|(X_2,X_3)}(x_1|(2,1)) = \begin{cases} \dfrac{x_1}{3}, \ x_1 \in \{1,2\}, \\ 0, \ \text{otherwise.} \end{cases}$$

Alternatively, since X_1, X_2 and X_3 are independent and X_1 and (X_2, X_3) are independent (why!), thus for fixed $(x_2,x_3) \in \mathbb{R}^2$ such that $f_{X_2,X_3}(x_2,x_3) > 0$,

$$f_{X_1|(X_2,X_3)}(x_1|(x_2,x_3)) = f_{X_1}(x_1) \ \forall\, x_1 \in \mathbb{R}$$
$$\implies f_{X_1|(X_2,X_3)}(x_1|(2,1)) = \begin{cases} \dfrac{x_1}{3}, \ x_1 \in \{1,2\}, \\ 0, \ \text{otherwise.} \end{cases}$$

(f) By (c), X_1 and X_3 are indepedent.

(g) $\Pr(X_1 = X_2 = X_3) = \displaystyle\sum_{\substack{\underline{x} \in S_X \\ x_1 = x_2 = x_3}} \frac{x_1 x_2 x_3}{72} = P(X_1 = X_2 = X_3 = 1) = \frac{1}{72}.$

16.0.3. Continuous Random Vectors

Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional random vector with d.f. F.

Definition 16.10. *The random vector $\underline{X}$ is called a continuous random vector if there exists a non-negative function $f : \mathbb{R}^p \to \mathbb{R}$ such that for any rectangale set A in $\mathbb{R}^p$*

$$P(\underline{X} \in A) = \int \cdots \int_A f(\underline{t}) \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_p, \quad \underline{t} = (t_1, t_2 \ldots, t_p).$$

The function $f(\cdot)$ is called probability density function of $\underline{X}$ and the set

$$S = \{\underline{x} \in \mathbb{R}^p : \Pr(x_i - h_i < x_i \leq x_i + h_i,\ i = 1, 2, \ldots, p) > 0\ \forall h_i > 0,\ \ i = 1, 2, \ldots, p\}$$

is called the support of F (or of $\underline{X}$).

Remark 16.11. *(a) In particular, if for fixed $\underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p$ if $A = (-\infty, x_1] \times (-\infty, x_2] \times \cdots \times (-\infty, x_p]$, then*

$$F(x_1, x_2, \ldots, x_p) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_p} f(t_1, t_2, \ldots, t_p) \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_p.$$

(b) If $\underline{X}$ is continuous random vector then its d.f. F is a continuous function.

(c) For a continuous random vector if its p.d.f. $f(\underline{x})$ is a piecewise continuous function then from the fundamental theorem of multivariable calculus

$$f(\underline{x}) = \frac{\partial^p}{\partial x_1 \partial x_2 \ldots \partial x_p} F(\underline{x}), \ \ \underline{x} \in \mathbb{R}^p,$$

whenever the derivative is defined.

(d) There are random vectors that are neither discrete nor continuous.

(e) If $\underline{X}$ is a continuous random vector with p.d.f. $f(\cdot)$ then $P(\underline{X} = \underline{a}) = \int \cdots \int_{\{\underline{x} = a\}} f(\underline{t}) \mathrm{d}\underline{t} = 0$.

(f) As in the univariate case the p.d.f. of a contnuous r.v. is not unique and it has different versions.

(g) It can be shown that if $\underline{X}$ is a p-dimensional random vector with continuous d.f. $F(\cdot)$ such that

$$\frac{\partial^p}{\partial x_1 \partial x_2 \ldots \partial x_p} F(x_1, x_2, \ldots, x_p)$$

exists everywhere except (possibly) on a set C comprising of countable number of curves (having 0 volume in $\mathbb{R}^p$) and

$$\int_{\mathbb{R}^p \setminus C} \frac{\partial^p}{\partial x_1 \partial x_2 \ldots \partial x_p} F(x_1, x_2, \ldots, x_p) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_p = 1.$$

Then $\underline{X}$ is a continuous random vector with p.d.f.

$$f(\underline{x}) = \begin{cases} \dfrac{\partial^p}{\partial x_1 \partial x_2 \ldots \partial x_p} F(x_1, x_2, \ldots, x_p), & \text{if } \underline{x} \in \mathbb{R}^p \setminus C, \\ 0, \quad \text{if } \underline{x} \in C. \end{cases}$$

(h) Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a continuous random vector with joint p.d.f. $f_{\underline{X}}(\underline{x})$ and d.f. $F_{\underline{X}}(\underline{x})$. Then for $q \in \{1, 2, \ldots, p-1\}$ and $\underline{x} = (x_1, x_2, \ldots, x_q) \in \mathbb{R}^q$, the marginal of $(X_1, X_2, \ldots, X_q)$ is a continuous random vector with p.d.f.

$$f_{X_1, X_2, \ldots, X_q}(x_1, x_2, \ldots, x_q) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}}(x_1, x_2, \ldots, x_q, t_{q+1}, \ldots, t_p) \mathrm{d}t_{q+1} \mathrm{d}t_{q+2} \ldots \mathrm{d}t_q.$$

Thus, marginal distribution of a continuous random vector $\underline{X}$ are continuous. Its p.d.f. is obtained by integrating the p.d.f. of $\underline{X}$ with respect to other (unwanted) variables .

Conditional distributions of continuous random vectors:

For simplicity consider $p = 2$ and let $\underline{X} = (X_1, X_2)$ be a random vector (discrete or continuous) with d.f. $F_{X_1,X_2}(x_1, x_2)$. Suppose that, for $x_1 \in S_{X_1}$ (the support of distribution of X_1) we want to define conditional d.f. of X_2 given $X_1 = x_1$. If X_1 is a continuous r.v. then $\Pr(X_1 = x_1) = 0$, $\forall x_1 \in \mathbb{R}$ and therefore $\Pr(X_2 \le x | X_1 = x_1)$ is not defined for any $x_1 \in \mathbb{R}$; although it is defined for discrete r.v. X_1 when $x_1 \in S_{X_1}$. Thus we define the conditional r.v. X_1 when $x_1 \in S_{X_1}$. Thus we define the conditional d.f. of X_2 given $X_1 = x_1$, thruogh the limiting argument

$$\begin{aligned} F_{X_2|X_1}(x, x_1) &= \lim_{h \downarrow 0} \Pr(X_2 \le x | x_1 - h < X_1 \le x_1) \\ &= \lim_{h \downarrow 0} \frac{\Pr(X_2 \le x | x_1 - h < X_1 \le x_1)}{\Pr(x_1 - h < X_1 \le x_1)} = \lim_{h \downarrow 0} \frac{F_{X_1,X_2}(x_1, x) - F_{X_1,X_2}(x_1 - h, x)}{F_{X_1}(x_1) - F_{X_1}(x_1 - h)}. \end{aligned}$$

Clearly if $\underline{X} = (X_1, X_2)$ is discrete and $x_1 \in S_{X_1}$, then

$$F_{X_2|X_1}(x, x_1) = \lim_{h \downarrow 0} \frac{F_{X_1,X_2}(x_1, x) - F_{X_1,X_2}(x_1 - h, x)}{F_{X_1}(x_1) - F_{X_1}(x_1 -)} = \frac{\Pr(X_1 = x_1, X_2 \le x)}{\Pr(X_1 = x_1)} = \Pr(X_2 \le x | X_1 = x_1).$$

Also, if $\underline{X} = (X_1, X_2)$ is continuous random vector with p.d.f. $f(x_1, x_2)$ hen

$$F_{X_2|X_1}(x, x_1) = \lim_{h \downarrow 0} \frac{\frac{1}{h} \int_{-\infty}^{x} \int_{x_1 - h}^{x_1} f_{X_1,X_2}(y_1, y_2) \mathrm{d}y_1 \mathrm{d}y_2}{\frac{F_{X_1}(x_1) - F_{X_1}(x_1 - h)}{h}} = \frac{\int_{-\infty}^{x} f_{X_1,X_2}(x_1, y_2) \mathrm{d}y_2}{f_{X_1}(x_1)}.$$

This implies conditional distribution of X_2 given $X_1 = x_1$ (provided $f_{X_1}(x_1) > 0$) is continuous with p.d.f.

$$f_{X_2|X_1}(x|x_1) = \frac{f_{X_1,X_2}(x_1, x) \mathrm{d}y_2}{f_{X_1}(x_1)},\ x \in \mathbb{R}$$

provided $f_{X_1}(x_1) > 0$.

The above discussion easily extends to general $p \ge 2$ by defining conditional d.f. of $\underline{X}_2 = (X_{q+1}, \ldots, X_p)$ given $\underline{X}_1 = (X_1, X_2, \ldots, X_q) = (x_1, x_2, \ldots, x_q) = \underline{x}_1$ as

$$F_{\underline{X}_2|\underline{X}_1}(\underline{x}_2|\underline{x}_1) = \lim_{h \downarrow 0} \Pr(X_q \le x_j,\ j = q+1, \ldots, p | x_i - h_i < X_i \le x_i, i = 1, 2, \ldots, q),$$

where $\underline{x}_2 = (x_{q+1}, x_{q+2}, \ldots, x_p) \in S_{\underline{X}_2}$.

Definition 16.12. *Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional random vector with joint p.d.f. $f_{\underline{X}}(\cdot)$. Let $q \in \{1, 2, \ldots, p-1\}$, $\underline{X}_1 = (X_1, X_2, \ldots, X_q)$ and $\underline{X}_2 = (X_{q+1}, \ldots, X_p)$. Then the conditional p.d.f. of $\underline{X}_2$ given $\underline{X}_1 = \underline{x}_1$ is defined by*

$$f_{\underline{X}_2|\underline{X}_1}(\underline{x}_2|\underline{x}_1) = \frac{f_{\underline{X}_1,\underline{X}_2}(\underline{x}_1, \underline{x}_2)}{f_{\underline{X}_1}(\underline{x}_1)} = \frac{f_{\underline{X}}(\underline{x}_1, \underline{x}_2)}{f_{\underline{X}_1}(\underline{x}_1)},\ \underline{x}_2 \in \mathbb{R}^{q-1}.$$

Theorem 16.13. *Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a continuous random vector with joint p.d.f. $f_{\underline{X}}(\cdot)$ and marginal p.d.f.s $f_{X_i}(\cdot)$, $i = 1, 2, \ldots, p$. Then $X_1, X_2, \ldots, X_p$ are independent iff*

$$f_{X_1,X_2,\ldots,X_p}(x_1, \ldots, x_p) = \prod_{i=1}^{p} f_{X_i}(x_i),\ \ \underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p.$$

Proof. Exercise. ☐

Remark 16.14. *(a) Let $S_{\underline{X}}$ be the support of distribution $\underline{X} = (X_1, X_2, \ldots, X_p)$ and let $S_{\underline{X_i}}$ be the support of distribution of X_i, $i = 1, 2, \ldots, p$. It can be shown that if $X_1, X_2, \ldots, X_p$ are indepedent then $S_{\underline{X}} = \prod_{i=1}^{p} S_{\underline{X_i}}$ (cartesian product).*

(b) Let $\underline{X} = (X_1, X_2)$ be a continuous random vector. Then X_1 and X_2 are indepedent iff, $\forall\, x_1 \in S_{X_1}$,

$$f_{X_2|X_1}(x_2|x_1) = f_{X_2}(x_2)\ \forall x_2 \in \mathbb{R}.$$

Theorem 16.15. *Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a continuous random vector with joint p.d.f. $f_{\underline{X}}(\cdot)$ and marginal p.d.f.s $f_{X_i}(\cdot)$, $i = 1, 2, \ldots, p$. Then $X_1, X_2, \ldots, X_p$ are independent iff*

$$f_{X_1, X_2, \ldots, X_p}(x_1, \ldots, x_p) = \prod_{i=1}^{p} g_i(x_i), \ \ \underline{x} \in \mathbb{R}^p,$$

for some non-negative functions $g_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, \ldots, p$. In that case $f_{X_i}(x) = c_i g_i(x)$, $x \in \mathbb{R}$, for some positive constants $c_i,\ i = 1, 2, \ldots, p$.