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

Lecture 18: Functions of Random Vector

Instructor: Dr. Kuldeep Kumar Kataria *Scribe:*

Theorem 18.1. *Suppose that the joint m.g.f.* $M_{\underline{X}}(\underline{t})$ *is finite on a rectangle* $(-\underline{a}, \underline{a}) \in \mathbb{R}^p$, $\underline{a} > 0$. *Then* $M_{\underline{X}}(\underline{t})$ *posseses partial derivatives of all order in* $(-\underline{a}, \underline{a})$. *Furthermore, for non-negative integers* $k_1, k_2, \ldots, k_p$

$$E\left(X_1^{k_1} X_2^{k_2} \ldots X_p^{k_p}\right) = \left[\frac{\partial^{k_1+k_2+\cdots+k_p}}{\partial t_1^{k_1} \partial t_2^{k_2} \ldots \partial t_p^{k_p}} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}}.$$

Proof. (*We give an outline of the proof*).

$$\begin{aligned} M_{\underline{X}}(t_1, t_2, \ldots, t_p) = E\left(e^{\sum_{i=1}^p t_i X_i}\right) &= \int_{\mathbb{R}^p} e^{\sum_{i=1}^p t_i x_i} f_{\underline{X}}(\underline{x}) \mathrm{d}\underline{x}, \\ \frac{\partial^{k_1+\cdots+k_p}}{\partial t_1^{k_1} \cdots \partial t_p^{k_p}} M_{\underline{X}}(\underline{t}) &= \int_{\mathbb{R}^p} x_1^{k_1} \cdots x_p^{k_p} e^{\sum_{i=1}^p t_i X_i} f_{\underline{X}}(\underline{x}) \mathrm{d}\underline{x}, \\ \left[\frac{\partial^{k_1+\cdots+k_p}}{\partial t_1^{k_1} \cdots \partial t_p^{k_p}} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}} &= \int_{\mathbb{R}^p} x_1^{k_1} \cdots x_p^{k_p} f_{\underline{X}}(\underline{x}) \mathrm{d}\underline{x} = E\left(X_1^{k_1} X_2^{k_2} \ldots X_p^{k_p}\right). \end{aligned}$$

This completes the proof. $\square$

Let $\psi_{\underline{X}}(\underline{t}) = \ln M_{\underline{X}}(\underline{t})$, $\underline{t} \in (-\underline{a}, \underline{a})$. Then

$$\begin{aligned} E(X_i) &= \left[\frac{\partial}{\partial t_i} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}} = \left[\frac{\partial}{\partial t_i} \psi_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}}, \\ E(X_i^m) &= \left[\frac{\partial^m}{\partial t_i^m} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}}, \ m = 1, 2, \ldots, \ i = 1, 2, \ldots, p, \\ \mathrm{Var}(X_i) &= \left[\frac{\partial^2}{\partial t_i^2} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}} - \left(\left[\frac{\partial}{\partial t_i} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}}\right)^2 = \left[\frac{\partial^2}{\partial t_i^2} \psi_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}}, \ i = 1, 2, \ldots, p, \end{aligned}$$

provided $M_{\underline{X}}(\underline{t})$ is finite on $(-\underline{a}, \underline{a})$, for some $\underline{a} > 0$. For $i \neq j$, if $M_{\underline{X}}(\underline{t})$ is finite on $(-\underline{a}, \underline{a})$, for some $\underline{a} > 0$,

$$\begin{aligned} \mathrm{Cov}(X_i, X_j) &= E(X_i X_j) - E(X_i)E(X_j) = E((X_i - E(X_i))(X_j - E(X_j))) \\ &= \left[\frac{\partial^2}{\partial t_i \partial t_j} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}} - \left[\frac{\partial}{\partial t_i} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}} \left[\frac{\partial}{\partial t_j} M_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}} = \left[\frac{\partial^2}{\partial t_i \partial t_j} \psi_{\underline{X}}(\underline{t})\right]_{\underline{t}=\underline{0}}. \end{aligned}$$

Moreover,

$$\begin{aligned} M_{\underline{X}}(0, \ldots, 0, t_i, 0, \ldots, 0) &= E(e^{t_i X_i}) = M_{X_i}(t_i), \ i = 1, 2, \ldots, p, \\ M_{\underline{X}}(0, \ldots, 0, t_i, 0, \ldots, 0, t_j, 0, \ldots, 0) &= E(e^{t_i X_i + t_j X_j}) = M_{X_i, X_j}(t_i, t_j), \end{aligned}$$

provided the m.g.f. is finite.

18.0.1. Equality in Distribution

Definition 18.2. *Two p-dimensional random vectors $\underline{X}$ and $\underline{Y}$ are said to have the same distribution (written as $\underline{X} \stackrel{d}{=} \underline{Y}$) if $F_{\underline{X}}(\underline{x}) = F_{\underline{Y}}(\underline{x}) \; \forall \; \underline{x} \in \mathbb{R}^p$.*

Theorem 18.3. *(a) Let $\underline{X}$ and $\underline{Y}$ be discrete random vectors with p.m.f.s $f_{\underline{X}}(\cdot)$ and $f_{\underline{Y}}(\cdot)$, respectively. Then*

$$\underline{X} \stackrel{d}{=} \underline{Y} \iff f_{\underline{X}}(\underline{x}) = f_{\underline{Y}}(\underline{x}) \; \forall \; \underline{x} \in \mathbb{R}^p.$$

(b) Let $\underline{X}$ and $\underline{Y}$ be continuous random vectors. Then

$$\underline{X} \stackrel{d}{=} \underline{Y} \iff f_{\underline{X}}(\underline{x}) = f_{\underline{Y}}(\underline{x}) \; \forall \; \underline{x} \in \mathbb{R}^p,$$

for some versions $f_{\underline{X}}(\cdot)$ and $f_{\underline{Y}}(\cdot)$ of p.d.f.s of $\underline{X}$ and $\underline{Y}$, respectively.

(c) Let $\underline{X}$ and $\underline{Y}$ be p-dimensional random vectors and let $\psi : \mathbb{R}^p \to \mathbb{R}^q$ be a given function. Then

$$\underline{X} \stackrel{d}{=} \underline{Y} \iff \psi(\underline{X}) \stackrel{d}{=} \psi(\underline{Y})$$

(d) Let $\underline{X}$ and $\underline{Y}$ be p-dimensional random vectors with finite m.g.f.s $M_{\underline{X}}(\underline{t})$ and $M_{\underline{Y}}(\underline{t})$ on a rectangle $(-\underline{a}, \underline{a})$, for some $\underline{a} > 0$. Then

$$M_{\underline{X}}(\underline{t}) = M_{\underline{Y}}(\underline{t}) \; \forall \; (-\underline{a}, \underline{a}) \implies \underline{X} \stackrel{d}{=} \underline{Y}.$$

18.0.2. Some Generalizations

Let $\underline{X}_i$: a p_i- dimensional random vector, $i = 1, 2, \ldots, m$. $F_{\underline{X}_i}$: d.f. of $\underline{X}_i$, $i = 1, 2, \ldots, m$, $f_{\underline{X}_i}$: p.m.f. / p.d.f. of $\underline{X}_i$, $i = 1, 2, \ldots, m$, $\sum_{i=1}^{p} p_i = p$, $\underline{X} = (\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_m)$: p-dimensional random vector with d.f. $F_{\underline{X}_i}(\cdot)$ and p.m.f. / p.d.f. $f_{\underline{X}_i}(\cdot)$.

Definition 18.4. *The random vectors $\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}$ are said to be independent if for any subcollection $\{\underline{X_{i_1}}, \underline{X_{i_2}}, \ldots, \underline{X_{i_q}}\}$ of $\{\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}\}$ $(2 \le q \le m)$*

$$F_{\underline{X_{i_1}}, \underline{X_{i_2}}, \ldots, \underline{X_{i_q}}}(\underline{x_1}, \underline{x_2}, \ldots, \underline{x_q}) = \prod_{j=1}^{q} F_{\underline{X_{i_j}}}(\underline{x_j}) \; \forall \; \underline{x} = (\underline{x_1}, \underline{x_2}, \ldots, \underline{x_q}) \in \mathbb{R}^{\sum_{j=1}^{q} p_{i_j}}.$$

Remark 18.5. *$\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}$ are independent $\implies$ random variables in any subset of $\{\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}\}$ are independent.*

Theorem 18.6. *(a) The following statements are equivalent:*

(i) $\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}$ are independent random vectors.

(ii) $F_{\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}}(\underline{x_1}, \underline{x_2}, \ldots, \underline{x_m}) = \prod_{i=1}^{m} F_{\underline{X_i}}(x_i) \; \forall \; \underline{x} = (\underline{x_1}, \underline{x_2}, \ldots, \underline{x_m}) \in \mathbb{R}^p$.

(iii) $f_{\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}}(\underline{x_1}, \underline{x_2}, \ldots, \underline{x_m}) = \prod_{i=1}^{m} f_{\underline{X_i}}(x_i) \; \forall \; \underline{x} = (\underline{x_1}, \underline{x_2}, \ldots, \underline{x_m}) \in \mathbb{R}^p$.

(iv) $f_{\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}}(\underline{x_1}, \underline{x_2}, \ldots, \underline{x_m}) = \prod_{i=1}^{m} g_i(x_i) \; \forall \; \underline{x} = (\underline{x_1}, \underline{x_2}, \ldots, \underline{x_m}) \in \mathbb{R}^p$ for some non-negative real valued function $g_i : \mathbb{R}^p \to \mathbb{R}$, $i = 1, 2, \ldots, m$.

(v) $\Pr(\underline{X_i} \in A_i, \; i =, 2, \ldots, m) = \prod_{i=1}^{m} \Pr(\underline{X_i} \in A_i) \; \forall \; A_i \in \mathscr{B}_{p_i}$, $i = 1, 2, \ldots, m$.

(b) If $\underline{X_1}, \underline{X_2}, \ldots, \underline{X_m}$ are independent random vectors, then

(i) $E\left(\prod_{i=1}^{m} \psi_i(\underline{X_i})\right) = \prod_{i=1}^{m} E\left(\psi_i(\underline{X_i})\right)$ for any functions ψ_i, $i = 1, 2, \ldots, m$.

(ii) $\psi_1(\underline{X_1}), \psi_2(\underline{X_2}), \ldots, \psi_m(\underline{X_m})$ are independent random vectors for any functions $\psi_1, \psi_2, \ldots, \psi_m$.

Definition 18.7. *Let Δ be an arbitrary index set. The random vectors $\{\underline{X}_\lambda : \lambda \in \Delta\}$ are said to be independent if random variables in any finite subcollection of $\{\underline{X}_\lambda : \lambda \in \Delta\}$ are independent.*

Theorem 18.8. *Under the notation of Theorem 18.6, $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_m$ are independent random vectors $\iff$ for some $\underline{a} > 0$ and $\forall\, \underline{t} = (\underline{t}_1, \underline{t}_2, \ldots, \underline{t}_m) \in (-\underline{a}, \underline{a})$,*

$$M_{\underline{X}}(\underline{t}_1, \underline{t}_2, \ldots, \underline{t}_m) = \prod_{i=1}^{m} M_{\underline{X}_i}(\underline{t}_i).$$

18.0.3. Functions of Random Vector

Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional random vector with p.m.f. / p.d.f. $f(\cdot)$. Let $g : \mathbb{R}^p \to \mathbb{R}^q$, where $1 \leq q \leq p$ be a function defined on $\mathbb{R}^p$ and taking values in $\mathbb{R}^q$. Sometimes it may be of interest to derive the probability distribution of $\underline{Y} = g(\underline{X})$.

Definition 18.9. *(a) Let $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ be a collection of iid random vectors each having the (joint) d.f. F and the same p.m.f. / p.d.f. $f(\cdot)$. We call $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ a random sample (r.s.) of size n from a distribution having d.f. F (p.m.f. / p.d.f. $f(\cdot)$). In other words a random sample is a collection of iid random vectors.*

(b) A function of one or more random vectors that does not depend on any unknown parameter is called a statistic.

Example 18.10. *Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution having p.d.f.*

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & x > 0, \\ 0, & \textit{otherwise}, \end{cases}$$

where $\theta \in \mathfrak{H} = (0, \infty)$ is unknown. Then $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is a statistic (called sample mean) but $X_1 - \theta$ is not a statistic. Some other statistic are:

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \text{ or } S_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \longrightarrow \textit{Sample Variance},$$

$$X_{r:n} = r\text{-th smallest of } X_1, X_2, \ldots, X_n,\ r = 1, 2, \ldots, n \text{ so that}$$
$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n} \longrightarrow r\text{-th order statistic},\ r = 1, \ldots, n,$$

$$X_{[np]:n}, 0 < p < 1;\ [x] = \text{largest integer} \leq x \longrightarrow p\text{-th sample quantile},$$

$$X_{[n/4]:n} \longrightarrow \text{sample lower quantile}, \quad X_{[3n/4]:n} \longrightarrow \text{sample upper quantile},$$

$$M = \begin{cases} X_{\frac{n+1}{2}:n}, & \text{if } n \text{ is odd}, \\ \frac{X_{\frac{n}{2}:n} + X_{\frac{n}{2}+1:n}}{2}, & \text{if } n \text{ is even}, \end{cases} \longrightarrow \text{sample median},$$

$$S_n = \sqrt{S_n^2} \text{ or } S_{n-1} = \sqrt{S_{n-1}^2} \longrightarrow \text{sample standard deviation},$$

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2\right)}} \longrightarrow \text{sample correlation coefficient}.$$

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution having d.f. F and p.m.f. / p.d.f. $f(\cdot)$. Then the joint d.f. of $\underline{X} = (X_1, X_2, \ldots, X_n)$ is

$$F_{\underline{X}}(\underline{x}) = \prod_{i=1}^{n} F(x_i), \quad \underline{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

and the joint p.m.f. / p.d.f. of $\underline{X}$ is

$$f_{\underline{X}}(\underline{x}) = \prod_{i=1}^{n} f(x_i), \quad \underline{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

Theorem 18.11. *If $X_1, X_2, \ldots, X_n$ is a random sample, then*

$$(X_1, X_2, \ldots, X_n) \stackrel{d}{=} (X_{\beta_1}, X_{\beta_2}, \ldots, X_{\beta_n})$$

for any permutation $(\beta_1, \beta_2, \ldots, \beta_n)$ of $(1, 2, \ldots, n)$.

Example 18.12. *Let $X_1, X_2, \ldots, X_n$ be a random sample from a given distribution.*

(a) If X_1 is a continuous r.v. then $\Pr(X_1 < X_2 < \cdots < X_n) = \Pr(X_{\beta_1} < X_{\beta_2} < \cdots < X_{\beta_n}) = \frac{1}{n!}$, for any permutation $(\beta_1, \beta_2, \ldots, \beta_n)$ of $(1, 2, \ldots, n)$.

(b) If X_1 is a continuous r.v. then for any $r \in \{1, 2, \ldots, n\}$, $\Pr(X_i = X_{r:n}) = \frac{1}{n}$, $i = 1, 2, \ldots, n$.

(c) $E\left(\frac{X_i}{X_1 + X_2 + \cdots + X_n}\right) = \frac{1}{n}$, $i = 1, 2, \ldots, n$.

(d) $E\left(X_i \middle| \sum_{j=1}^{n} X_j = t\right) = \frac{t}{n}$, $i = 1, 2, \ldots, n$.

Solution (a)

$$\begin{aligned}
& X_1 \text{ is a continuous r.v.} \implies \underline{X} = (X_1, X_2, \ldots, X_n) \text{ is a continuous random vector. (Why?)} \\
\implies & (X_1, X_2, \ldots, X_n) \stackrel{d}{=} (X_{\beta_1}, X_{\beta_2}, \ldots, X_{\beta_n}) \text{ for any permutation } (\beta_1, \beta_2, \ldots, \beta_n) \text{ of } (1, 2, \ldots, n) \text{ and} \\
& \Pr(\text{all } X_i\text{'s are distinct}) = 1 \\
\implies & (X_1, X_2, \ldots, X_n) \stackrel{d}{=} (X_{\beta_1}, X_{\beta_2}, \ldots, X_{\beta_n}) \text{ for any permutation } (\beta_1, \beta_2, \ldots, \beta_n) \text{ of } (1, 2, \ldots, n) \text{ and} \\
& \sum_{\beta \in S_n} \Pr(X_{\beta_1} < X_{\beta_2} < \cdots < X_{\beta_n}) = 1, \text{ where } S_n \text{ is the set of all permutation of } (1, 2, \ldots, n) \\
\implies & \Pr(X_{\beta_1} < X_{\beta_2} < \cdots < X_{\beta_n}) = \Pr(X_1 < X_2 < \cdots < X_n) = \frac{1}{n!}.
\end{aligned}$$

(b)

$$\begin{aligned}
& \text{For any } i = 1, 2, \ldots, n, (X_1, X_2, \ldots, X_i, \ldots, X_n) \stackrel{d}{=} (X_i, X_2, \ldots, X_1, \ldots, X_n) \\
\implies & X_{r:n},\ r\text{-th smallest of } (X_1, X_2, \ldots, X_i, \ldots, X_n) = r\text{-th smallest of } (X_i, X_2, \ldots, X_1, \ldots, X_n) \text{ and} \\
& \Pr(X_1 = r\text{-th smallest of } (X_1, X_2, \ldots, X_i, \ldots, X_n)) = \Pr(X_i = r\text{-th smallest of } (X_i, X_2, \ldots, X_1, \ldots, X_n)) \\
\implies & \Pr(X_1 = X_{r:n}) = \Pr(X_i = X_{r:n}), \ i = 1, 2, \ldots, n \\
& \text{since } \Pr(X_{1:n} < X_{2:n} < \cdots < X_{n:n}) = 1, \ \text{(by (a)), we have } \sum_{i=1}^{n} \Pr(X_i = X_{r:n}) = 1 \\
\implies & \Pr(X_1 = X_{r:n}) = \Pr(X_i = X_{r:n}) = \frac{1}{n}.
\end{aligned}$$

(c)

$$
\begin{aligned}
& (X_1, X_2, \ldots, X_i, \ldots, X_n) \stackrel{d}{=} (X_i, X_2, \ldots, X_1, \ldots, X_n) \\
\Longrightarrow & E\left(\frac{X_1}{X_1 + X_2 + \cdots + X_i + \cdots + X_n}\right) = E\left(\frac{X_i}{X_i + X_2 + \cdots + X_1 + \cdots + X_n}\right) \\
\Longrightarrow & E\left(\frac{X_1}{\sum_{j=1}^n X_j}\right) = E\left(\frac{X_i}{\sum_{j=1}^n X_j}\right) \text{ but } \sum_{i=1}^n E\left(\frac{X_i}{\sum_{j=1}^n X_j}\right) = E\left(\frac{\sum_{i=1}^n X_i}{\sum_{j=1}^n X_j}\right) = 1 \\
\Longrightarrow & E\left(\frac{X_i}{\sum_{j=1}^n X_j}\right) = E\left(\frac{X_1}{\sum_{j=1}^n X_j}\right) = \frac{1}{n}, \;\; i = 1, 2, \ldots, n.
\end{aligned}
$$

(d)

$$
\begin{aligned}
& (X_1, X_2, \ldots, X_i, \ldots, X_n) \stackrel{d}{=} (X_i, X_2, \ldots, X_1, \ldots, X_n) \\
\Longrightarrow & E\left(X_1 | X_1 + X_2 + \cdots + X_i + \cdots + X_n = t\right) = E\left(X_i | X_i + X_2 + \cdots + X_1 + \cdots + X_n = t\right) \\
\Longrightarrow & E\left(X_1 \middle| \sum_{j=1}^n X_j = t\right) = E\left(X_i \middle| \sum_{j=1}^n X_j = t\right) \text{ but } \sum_{i=1}^n E\left(X_i \middle| \sum_{j=1}^n X_j = t\right) = E\left(\sum_{i=1}^n X_i \middle| \sum_{j=1}^n X_j = t\right) = t.
\end{aligned}
$$

Therefore

$$
E\left(X_i \middle| \sum_{j=1}^n X_j = t\right) = E\left(X_1 \middle| \sum_{j=1}^n X_j\right) = \frac{t}{n}, \;\; i = 1, 2, \ldots, n.
$$

18.0.4. Distribution Function Technique

Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional random vector with d.f. F and p.m.f. / p.d.f. $f(\cdot)$. Also, let $\underline{g} : \mathbb{R}^p \to \mathbb{R}^q : \underline{g} = (g_1, g_2, \ldots, g_q)$, $\underline{Y} = (Y_1, Y_2, \ldots, Y_q) = (g_1(\underline{X}), g_2(\underline{X}), \ldots, g_q(\underline{X}))$. We are interested in the distribution of random vector $\underline{Y}$.

One can first find the d.f. of $\underline{Y} = (Y_1, Y_2, \ldots, Y_q)$

$$
F_{\underline{Y}}(y_1, y_2, \ldots, y_q) = \Pr(g_1(\underline{X}) \leq y_1, g_2(\underline{X}) \leq y_2, \ldots, g_q(\underline{X}) \leq y_q), \;\; \underline{y} = (y_1, y_2, \ldots, y_q) \in \mathbb{R}^q,
$$

and then find the p.m.f. / p.d.f. of $\underline{Y} = (Y_1, Y_2, \ldots, Y_q)$.

Example 18.13. *Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution having d.f. F, p.m.f. / p.d.f. f and support S. Let $Y_1 = \min\{X_1, X_2, \ldots, X_n\}$ and $Y_2 = \max\{X_1, X_2, \ldots, X_n\}$.*

(a) Find the joint d.f. of $\underline{Y} = (Y_1, Y_2)$.

(b) Find the marginal d.f.s of Y_1 and Y_2 using findings of (a).

(c) Find the marginal d.f.s of Y_1 and Y_2 directly (that is, without using (a)).

(d) Find the marginal p.m.f. / p.d.f. $\underline{Y} = (Y_1, Y_2)$ using findings in (b).

Solution: (a) For $(y_1, y_2) \in \mathbb{R}^2$,

$$\begin{aligned}
F_{\underline{Y}}(y_1, y_2) &= \Pr(Y_1 \le y_1, Y_2 \le y_2) \\
&= \Pr(\min\{X_1, X_2, \ldots, X_n\} \le y_1, \max\{X_1, X_2, \ldots, X_n\} \le y_2) \\
&= \Pr(\max\{X_1, X_2, \ldots, X_n\} \le y_2) - \Pr(\min\{X_1, X_2, \ldots, X_n\} > y_1, \max\{X_1, X_2, \ldots, X_n\} \le y_2) \\
&= \Pr(X_i \le y_2,\ i = 1, 2, \ldots, n) - \Pr(X_i > y_1,\ i = 1, 2, \ldots, n, X_i \le y_2, i = 1, 2, \ldots, n) \\
&= \prod_{i=1}^{n} \Pr(X_i \le y_2) - \Pr(y_1 < X_i \le y_2, i = 1, 2, \ldots, n) \\
&= \prod_{i=1}^{n} \Pr(X_i \le y_2) - \prod_{i=1}^{n} \Pr(y_1 < X_i \le y_2) = \begin{cases} [F(y_2)]^n - [F(y_2) - F(y_1)]^n, & -\infty < y_1 < y_2 < \infty, \\ [F(y_2)]^n, & -\infty < y_2 < y_1 < \infty. \end{cases}
\end{aligned}$$

(b)

$$F_{Y_1}(y_1) = \lim_{y_2 \to \infty} F_{\underline{Y}}(y_1, y_2) = 1 - [1 - F(y_1)]^n, -\infty < y_1 < \infty,$$
$$F_{Y_2}(y_2) = \lim_{y_1 \to \infty} F_{\underline{Y}}(y_1, y_2) = [F(y_2)]^n, -\infty < y_2 < \infty.$$

(c)

$$\begin{aligned}
F_{Y_1}(y_1) &= \Pr(Y_1 \le y_1) \\
&= \Pr(\min\{X_1, X_2, \ldots, X_n\} \le y_1) \\
&= 1 - \Pr(\min\{X_1, X_2, \ldots, X_n\} > y_1) \\
&= 1 - \Pr(X_i > y_1, i = 1, 2, \ldots, n) = 1 - \prod_{i=1}^{n} \Pr(X_i > y_1) = 1 - [1 - F(y_1)]^n, -\infty < y_1 < \infty.
\end{aligned}$$

$$\begin{aligned}
F_{Y_2}(y_2) &= \Pr(Y_2 \le y_2) \\
&= \Pr(\max\{X_1, X_2, \ldots, X_n\} \le y_2) \\
&= \Pr(X_i \le y_2, i = 1, 2, \ldots, n) = \prod_{i=1}^{n} \Pr(X_i \le y_2) = [F(y_2)]^n, -\infty < y_2 < \infty.
\end{aligned}$$

(d) **Case I:** X_1 is a discrete r.v. Then $S_{X_1} = S_{Y_1} = S_{Y_2}$. For $y_1 \in S_{X_1}$

$$f_{Y_1}(y_1) = \Pr(Y_1 = y_1) = F_{Y_1}(y_1) - F_{Y_1}(y_1-) = [1 - F(y_1-)]^n - [1 - F(y_1)]^n.$$

Thus,

$$f_{Y_1}(y_1) = \begin{cases} [1 - F(y_1-)]^n - [1 - F(y_1)]^n, & \text{if } y_1 \in S_{X_1}, \\ 0, & \text{otherwise.} \end{cases}$$

Similarly,

$$f_{Y_2}(y_2) = F_{Y_2}(y_2) - F_{Y_2}(y_2-) = \begin{cases} [F(y_2)]^n - [F(y_2-)]^n, & \text{if } y_2 \in S_{X_1} \\ 0, & \text{otherwise} \end{cases}$$

Case II: X_1 is a continuous r.v.

Let $F(\cdot)$ be differentiable everywhere (except possibly on a set having length zero (that is, it does not contain any open interval)

$$f_{Y_1}(y) = \frac{\mathrm{d}}{\mathrm{d}y} (1 - [1 - F(y)]^n) = n [1 - F(y)]^{n-1} f(y), -\infty < y < \infty,$$
$$f_{Y_2}(y) = \frac{\mathrm{d}}{\mathrm{d}y} [F(y)]^n = n [F(y)]^{n-1} f(y), -\infty < y < \infty.$$

Example 18.14. *Let X_1 and X_2 be iid r.v.s with common p.d.f.*

$$f(x) = \begin{cases} 2x, \ 0 < x < 1, \\ 0, \ \textit{otherwise}. \end{cases}$$

Find the d.f. of $Y = X_1 + X_2$. Hence find the p.d.f. of Y.

Solution: The joint p.d.f. of $\underline{X} = (X_1, X_2)$ is

$$f_{\underline{X}}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = f(x_1) f(x_2) = \begin{cases} 4x_1x_2, \ 0 < x_1 < 1, \ 0 < x_2 < 1, \\ 0, \ \text{otherwise.} \end{cases}$$

For $y \in \mathbb{R}$,

$$F_Y(y) = \Pr(Y \le y) = \Pr(X_1 + X_2 \le y) = \underset{x_1+x_2 \le y}{\int_0^1 \int_0^1} 4x_1x_2 \mathrm{d}x_1 \mathrm{d}x_2.$$

Clearly for $y < 0$, $F_Y(y) = 0$ and for $y \ge 2$, $F_Y(y) = 1$. Now consider $y \in [0, 1)$,

$$F_Y(y) = \int_0^y \int_0^{y-x_1} 4x_1x_2 \mathrm{d}x_2 \mathrm{d}y_1 = \frac{y^4}{6}.$$

For $y \in [1, 2)$,

$$F_Y(y) = \int_0^{y-1} \int_0^1 4x_1x_2 \mathrm{d}x_2 \mathrm{d}y_1 + \int_{y-1}^1 \int_0^{y-x_1} 4x_1x_2 \mathrm{d}x_2 \mathrm{d}x_1 = (y-1)^2 + \frac{(4y-3) - (y+3)(y-1)^3}{6}.$$

Thus,

$$F_Y(y) = \begin{cases} 0, \ y < 0, \\ \frac{y^4}{6}, \ 0 \le y < 1, \\ (y-1)^2 + \frac{(4y-3)-(y+3)(y-1)^3}{6}, \ 1 \le y < 2, \\ 1, \ y \ge 2. \end{cases}$$

Clearly, Y is continuous r.v. with p.d.f.

$$f_Y(y) = \begin{cases} \frac{2}{3}y^3, \ 0 < y < 1, \\ 2(y-1) + \frac{2}{3[1-(y+2)(y-1)^2]}, \ 1 < y < 2, \\ 0, \ \text{otherwise.} \end{cases}$$