Lecture 15: Random Vectors and their Distribution Functions

Instructor: Dr. Kuldeep Kumar Kataria *Scribe:*

Let $X \sim N(\mu, \sigma^2)$. Then m.g.f. of X is

$$\begin{aligned}
M_X(t) = E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x, \text{ take } \frac{x-\mu}{\sigma} = z,\ x = (\mu + \sigma z) \\
&= \int_{-\infty}^{\infty} e^{t(\mu+\sigma z)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \mathrm{d}z \\
&= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2\sigma t z + \sigma^2 t^2)} \mathrm{d}z \\
&= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\sigma t)^2} \mathrm{d}z = e^{\mu t + \frac{1}{2}\sigma^2 t^2}, \quad t \in \mathbb{R}.
\end{aligned}$$

· Let $X \sim N(\mu, \sigma^2)$. Then $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$.

$$M_Z(t) = E(e^{t\frac{(X-\mu)}{\sigma}}) = e^{-\mu t/\sigma} M_X(t/\sigma) = e^{-\mu t/\sigma} e^{\mu t/\sigma + t^2/2} = e^{t^2/2}\ \forall\, t \in \mathbb{R} \rightarrow \text{m.g.f. of } N(0,1)$$
$$\implies Z \sim N(0,1).$$

Theorem 15.1. *Let $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0,1)$.*

(a) For $a \neq 0$, $b \in \mathbb{R}$, $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

(b) $Z \stackrel{d}{=} \frac{X-\mu}{\sigma} \sim N(0,1)$.

(c)

$$E(Z^r) = \begin{cases} 0, & \text{if } r = 1, 3, 5, \ldots, \\ \frac{r!}{2^{r/2}(r/2)!}, & \text{if } r = 2, 4, 6, \ldots. \end{cases}$$

(d) Mean = $\mu_1' = E(X) = \mu$; Variance = $\mu_2 = \sigma^2$; coefficient of skewness = $\beta_1 = 0$; kurtosis = $\nu_1 = 3$.

(e) $Z^2 \sim \chi_1^2$.

Proof. (a) Note that

$$\begin{aligned}
M_Y(t) = E(e^{tY}) &= E\left(e^{t(aX+b)}\right) \\
&= e^{bt} E\left(e^{(ta)X}\right) = e^{bt} M_X(at) \\
&= e^{bt} e^{\mu(at) + \frac{1}{2}\sigma^2 (at)^2} \\
&= e^{(a\mu+b)t + \frac{1}{2}(a^2\sigma^2)t^2},\ t \in \mathbb{R} \implies Y \sim N(a\mu + b, a^2\sigma^2).
\end{aligned}$$

(b) Follows from (a) by taking $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$.

(c) $M_Z(t) = e^{t^2/2} = \sum_{k=0}^{\infty} \frac{t^{2k}}{2^k k!}$, $t \in \mathbb{R}$.

$$E(Z^r) = \text{ Coefficient of } \frac{t^r}{r!} \text{ in the expansion of } M_Z(t) = \begin{cases} 0, & \text{if } r = 1, 3, 5, \ldots, \\ \frac{r!}{2^{r/2}(r/2)!}, & \text{if } r = 2, 4, 6, \ldots. \end{cases}$$

(d) $Z \stackrel{d}{=} \frac{X-\mu}{\sigma}$.

$$E\left(\frac{X-\mu}{\sigma}\right) = E(Z) = 0 \implies \mu_1' = E(X) = \mu,$$
$$E\left(\left(\frac{X-\mu}{\sigma}\right)^2\right) = E(Z^2) = 1 \implies \mu_2 = E((X-\mu)^2) = \sigma^2,$$
$$E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right) = E(Z^3) = 0 \implies \mu_3 = E((X-\mu)^3) = 0,$$
$$E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right) = E(Z^4) = 3 \implies \mu_4 = 3\sigma^4 = 3,$$
$$\text{Coefficient of skewness} = \beta_1 = \frac{\mu_3}{\mu_2^2} = 0, \;\; \text{kurtosis} = \frac{\mu_4}{\mu_2^2} = 3.$$

(e) Let $Y = Z^2$. Then

$$M_Y(t) = E(e^{tZ^2}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz^2} e^{-z^2/2} \mathrm{d}z = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(1-2t)}{2} z^2} \mathrm{d}z = (1-2t)^{-1/2}, \; t < \frac{1}{2} \implies Z^2 \sim \chi_1^2.$$

This completes the proof. □

Corollary 15.2. *Let $X_1, X_2, \ldots, X_k$ be independent and let $X_i \sim N(\mu_i, \sigma_i^2)$, $-\infty < \mu_i < \infty$, $\sigma_i > 0$, $i = 1, 2, \ldots, k$. Then $\sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi_k^2$.*

Remark 15.3. *(i) In $N(\mu, \sigma^2)$ distribution the parameters $\mu \in (-\infty, \infty)$ and $\sigma^2 > 0$ are respectively, the mean and variance of the distribution.*

(ii) If $X \sim N(\mu, \sigma^2)$, then

$$P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right), \;\; x \in \mathbb{R}.$$

Let τ_α be the $(1-\alpha)$th quantile of Φ then $\Phi(-\tau_\alpha) = 1 - \Phi(\tau_\alpha) = \alpha$. Tables for values of $\Phi(x)$ for different values of x are available in various text books.

Example 15.4. *Let $X \sim N(2, 4)$. Find $P(X \le 0)$, $P(|X| \ge 2)$, $P(1 < X \le 3)$ and $P(X \le 3 | X > 1)$.*

Solution: $P(X \le 0) = \Phi\left(\frac{0-2}{2}\right) = \Phi(-1) = 1 - \Phi(1) = .1587$,

$$P(|X| \ge 2) = P(X \le -2) + P(X \ge 2) = \Phi\left(\frac{-2-2}{2}\right) + 1 - \Phi\left(\frac{2-2}{2}\right)$$
$$= \Phi(-2) + 1 - \Phi(0) = 0.0228 + 0.5 = 0.5228,$$

$$P(1 < X \le 3) = P(X \le 3) - P(X \ge 1) = \Phi\left(\frac{3-2}{2}\right) + 1 - \Phi\left(\frac{1-2}{2}\right) = 2\Phi(0.5) - 1 = 2 \times 0.6915 - 1 = 0.383,$$

$$P(X \le 3 | X > 1) = \frac{P(1 < X \le 3)}{P(X > 1)} = \frac{.383}{1 - \Phi\left(\frac{1-2}{2}\right)} = \frac{0.383}{\Phi(0.5)} = 0.55599.$$

Theorem 15.5. *Let $X_1, X_2, \ldots, X_k$ be independent r.v.'s and let $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, k$. Let $a_1, a_2, \ldots, a_k$ be real constants such that $\sum_{i=1}^k a_i^2 > 0$. Then $Y = \sum_{i=1}^k a_i X_i \sim N(\sum_{i=1}^k a_i\mu_i, \sum_{i=1}^k a_i^2\sigma_i^2)$.*

Proof. Note that

$$\begin{aligned}
M_Y(t) = E(e^{t\sum_{i=1}^k a_i X_i}) = E\left(\prod_{i=1}^k e^{ta_iX_i}\right) &= \prod_{i=1}^k E(e^{ta_iX_i}), \quad \text{(independent of } X_i\text{'s)} \\
&= \prod_{i=1}^k M_{X_i}(ta_i) = \prod_{i=1}^k e^{\mu_i t a_i + \frac{1}{2}\sigma_i^2 t^2 a_i^2} = e^{(\sum_{i=1}^k a_i\mu_i)t + \frac{(\sum_{i=1}^k a_i^2\sigma_i^2)t}{2}} \\
&\to \text{m.g.f. of } N\left(\sum_{i=1}^k a_i\mu_i, \sum_{i=1}^k a_i^2\sigma_i^2\right).
\end{aligned}$$

By uniqueness of m.g.f.'s $Y \sim N(\sum_{i=1}^k a_i\mu_i, \sum_{i=1}^k a_i^2\sigma_i^2)$. □

15.1. Random Vectors and their Distribution Functions

Let $(\Omega, \mathcal{F}, P)$ be a given probability space. In many situations we may be interested in simultaneously studying two or more numerical characteristics of outcomes of a random experiment. This amounts to define a function

$$\underline{X} = (X_1, X_2, \ldots, X_p) : \Omega \to \mathbb{R}^p.$$

Example 15.6. *A fair coin is tossed three times independently. Then*

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \text{ and } P(\{\omega\}) = \frac{1}{8}\ \forall\, \omega \in \Omega.$$

Suppose that we are simultaneously interested in:

· number of heads in three tosses,

· number of heads in first two tosses.

Here we are interested in the function $(X, Y) : \Omega \to \mathbb{R}^2$ defined by

$$(X(\omega), Y(\omega)) = \begin{cases} (0,0) & \text{if } \omega = TTT, \\ (1,0) & \text{if } \omega = TTH, \\ (1,1) & \text{if } \omega = HTT, THT, \\ (2,1) & \text{if } \omega = HTH, THH, \\ (2,2) & \text{if } \omega = HHT, \\ (3,2) & \text{if } \omega = HHH. \end{cases}$$

The values assumed by (X, Y) are random with

$$\Pr\{(X,Y) = (x,y)\} = \begin{cases} \frac{1}{8}, & \text{if } (x,y) \in \{(0,0),(1,0),(2,2),(3,2)\}, \\ \frac{1}{4}, & \text{if } (x,y) \in \{(1,1),(2,1)\}, \\ 0, & \text{otherwise.} \end{cases}$$

Here, $\Pr((X,Y) \in \{(0,0),(1,0),(2,2),(3,2),(1,1),(2,1)\}) = 1$.

Definition 15.7. *Let $(\Omega, \mathcal{F}, P)$ be a given probability space. A function $\underline{X} = (X_1, X_2, \ldots, X_p) : \Omega \to \mathbb{R}^p$ (defined on the sample space Ω) is called a random vector (p-dimensional random vector). A one dimensional random vector is simply called a random variable.*

For any function $\underline{Y} = (Y_1, Y_2, \ldots, Y_p) : \Omega \to \mathbb{R}^p$ and $A \subseteq \mathbb{R}^p$, define $\underline{Y}^{-1} = \{\omega \in \Omega : \underline{Y}(\omega) \in A\}$. For probability space $(\Omega, \mathcal{F}, P)$ and a p-dimensional random vector $\underline{X} = (X_1, X_2, \ldots, X_p) : \Omega \to \mathbb{R}^p$, define $P_{\underline{X}}(B) = P(\underline{X}^{-1}(B))$, $B \in \mathscr{B}_p$ where for all practical purpose we take $\mathscr{B}_p$ to be power set of $\mathbb{R}^p$. We will simply write

$$P_{\underline{X}}(B) = P(\{\omega \in \Omega : \underline{X}(\omega) \in B\}) = \Pr(X \in B), \;\; B \in \mathscr{B}_p.$$

The following scenario has emerged: $(\Omega, \mathcal{F}, P) \xrightarrow{X} (\mathbb{R}^p, \mathscr{B}_p, P_{\underline{X}})$.

Theorem 15.8. *$(\mathbb{R}^p, \mathscr{B}_p, P_{\underline{X}})$ defined above is a probability space, i.e. $P_{\underline{X}}(\cdot)$ is a probability function defined on $\mathscr{B}_p$.*

Proof. Similar to the proof of random variable case. □

Definition 15.9. *The probability function $P_{\underline{X}}(\cdot)$ defined above is called the probability function / measure induced by random vector $\underline{X}$ and $(\mathbb{R}^p, \mathscr{B}_p, P_{\underline{X}})$ is called the probability space induced by random vector $\underline{X}$.*

The induced probability measure $P_{\underline{X}}(\cdot)$ describes the random behaviour of $\underline{X}$.

Example 15.10. *Consider the sample space defined in Example 15.6, where*

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \text{ and } P(\{\omega\}) = \frac{1}{8} \; \forall \, \omega \in \Omega$$

and $(X, Y) : \Omega \to \mathbb{R}^2$ is defined by

$$(X(\omega), Y(\omega)) = \begin{cases} (0,0) & \text{if } \omega = TTT, \\ (1,0) & \text{if } \omega = TTH, \\ (1,1) & \text{if } \omega = HTT, THT, \\ (2,1) & \text{if } \omega = HTH, THH, \\ (2,2) & \text{if } \omega = HHT, \\ (3,2) & \text{if } \omega = HHH. \end{cases}$$

Here, $(X, Y) : \Omega \to \mathbb{R}^2$ is a random vector with induced probability space $(\mathbb{R}^2, \mathscr{B}_2, P_{\underline{X}})$, where

$$P_{\underline{X}}(\{(i,j)\}) = \begin{cases} \frac{1}{8}, & \text{if } (i,j) \in \{(0,0), (1,0), (2,2), (3,2)\}, \\ \frac{1}{4}, & \text{if } (i,j) \in \{(1,1), (2,1)\}, \\ 0, & \text{otherwise}, \end{cases}$$

and for any $B \in \mathscr{B}_2$

$$P_X(B) = \sum_{(i,j) \in B \cap S} P_X(\{(i,j)\}), \;\; \text{where } S = \{(0,0), (1,0), (2,2), (3,2), (1,1), (2,1)\}.$$

Definition 15.11. *(a) The joint distribution function of a p-dimensional random vector $\underline{X} = (X_1, X_2, \ldots, X_p)$ is defined as*

$$F_{\underline{X}}(x_1, x_2, \ldots, x_p) = \Pr(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p), \;\; \underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p.$$

(b) The joint d.f. of any subset of random vectors $(X_1, X_2, \ldots, X_p)$ is called a marginal distribution function of $F_{\underline{X}}(\cdot)$ (or $\underline{X} = (X_1, X_2, \ldots, X_p)$).

Example 15.12. *$F_{X_1,X_2}(x,y)$, $(x,y)\in\mathbb{R}^2$, $F_{X_2}(x)$, $x\in\mathbb{R}$ and $F_{X_1,X_2,X_3}(x,y,z)$, $(x,y,z)\in\mathbb{R}^3$ are marginal d.f.s of $F_{X_1,X_2,X_3,X_4}(x_1,x_2,x_3,x_4)$, $(x_1,x_2,x_3,x_4)\in\mathbb{R}^4$.*

In the sequel we will describe a notation for writing down all the vertices of a p-dimensional rectangle in a compact form.

For $-\infty\le a_i<b_i<\infty,\ \ i=1,2,\ \ \underline{a}=(a_1,a_2)$ and $\underline{b}=(b_1,b_2)$, the vertices of two dimensional rectangle

$$(\underline{a},\underline{b}]=(a_1,b_1]\times(a_2,b_2]=\{(x,y)\in\mathbb{R}^2: a_1<x\le b_1,\ a_2<y\le b_2\}$$

are

$$\{(b_1,b_2),(a_1,b_2),(b_1,a_2),(a_1,a_2)\}=\{(b_1,b_2)\}\cup\{(a_1,b_2),(b_1,a_2)\}\cup\{(a_1,a_2)\}=\Delta_{0,2}\cup\Delta_{1,2}\cup\Delta_{2,2},\ \textit{say}.$$

Similarly, for $-\infty\le a_i<b_i<\infty,\ \ i=1,2,3,\ \ \underline{a}=(a_1,a_2,a_3)$ and $\underline{b}=(b_1,b_2,b_3)$, the vertices of three dimensional rectangle

$$(\underline{a},\underline{b}]=(a_1,b_1]\times(a_2,b_2]\times(a_3,b_3]=\{(x_1,x_2,x_3)\in\mathbb{R}^3: a_i<x_i\le b_i,\ i=1,2,3\}$$

are

$$\begin{aligned}&\{(b_1,b_2,b_3),(a_1,b_2,b_3),(b_1,a_2,b_3),(b_1,b_2,a_3),(a_1,a_2,b_3),(a_1,b_2,a_3),(b_1,a_2,a_3),(a_1,a_2,a_3)\}\\&\quad=\{(b_1,b_2,b_3)\}\cup\{(a_1,b_2,b_3),(b_1,a_2,b_3),(b_1,b_2,a_3)\}\cup\{(a_1,a_2,b_3),(a_1,b_2,a_3),(b_1,a_2,a_3)\}\cup\{(a_1,a_2,a_3)\}\\&\quad=\Delta_{0,3}\cup\Delta_{1,3}\cup\Delta_{2,3}\cup\Delta_{3,3},\ \textit{say}.\end{aligned}$$

In general, for $-\infty\le a_i<b_i<\infty,\ \ i=1,2,\dots,p,\ \ \underline{a}=(a_1,a_2,\dots,a_p)$ and $\underline{b}=(b_1,b_2,\dots,b_p)$ define

$$\Delta_{k,p}\equiv\Delta_{k,p}((\underline{a},\underline{b}])=\{\underline{z}\in\mathbb{R}^p: z_i\in\{a_i,b_i\},\ i=1,2,\dots,p \text{ and exactly } k \text{ of } z_i's \text{ are } a_j's\}\ \left(\rightarrow \text{has } \binom{p}{k} \text{ elements}\right)$$

where $(\underline{a},\underline{b}]=(a_1,b_1]\times(a_2,b_2]\times\cdots\times(a_p,b_p]$.

Then $\bigcup_{k=0}^{p}\Delta_{k,p}$ is the set of $2^p\ \left(=\sum_{k=0}^{p}\binom{p}{k}\right)$ vectors of p-dimensional rectangle $(\underline{a},\underline{b}]$.

Theorem 15.13. *For constants $-\infty\le a_i<b_i<\infty,\ \ i=1,2,\dots,p$*

$$\Pr(a_i<X_i\le b_i,\ i=1,2,\dots,p)=\sum_{k=0}^{p}(-1)^k\sum_{\underline{z}\in\Delta_{k,p}((\underline{a},\underline{b}])}F_{\underline{X}}(\underline{z}).$$

Proof. **Special cases:**

Case I: $p=1$

We have $\Delta_{0,1}((a_1,b_1])=\{b_1\}$ and $\Delta_{1,1}((a_1,b_1])=\{a_1\}$. Then

$$R.H.S.=F_{X_1}(b_1)-F_{X_1}(a_1)=\Pr(a_1<X_1\le b_1)=L.H.S.$$

Case II: $p=2$

Here $\Delta_{0,2}=\{(b_1,b_2)\}$, $\Delta_{1,2}=\{(a_1,b_2),(b_1,a_2)\}$ and $\Delta_{2,2}=\{(a_1,a_2)\}$. Thus

$$\begin{aligned}R.H.S.&=F_{\underline{X}}(b_1,b_2)-F_{\underline{X}}(a_1,b_2)-F_{\underline{X}}(b_1,a_2)+F_{\underline{X}}(a_1,a_2)\\&=\Pr(X_1\le b_1,\ X_2\le b_2)-\Pr(X_1\le a_1,\ X_2\le b_2)-\Pr(X_1\le b_1,\ X_2\le a_2)+\Pr(X_1\le a_1,\ X_2\le a_2)\\&=\Pr(a_1<X_1\le b_1,\ X_2\le b_2)-\Pr(a_1<X_1\le b_1,\ X_2\le a_2)\\&=\Pr(a_1<X_1\le b_1,\ a_2<X_2\le b_2)=L.H.S.\end{aligned}$$

Case III: $p = 3$

$$
\begin{aligned}
&\Pr(a_1 < X_1 \le b_1, a_2 < X_2 \le b_2, a_3 < X_3 \le b_3) \\
&= \Pr(a_1 < X_1 \le b_1, a_2 < X_2 \le b_2, X_3 \le b_3) - \Pr(a_1 < X_1 \le b_1, a_2 < X_2 \le b_2, X_3 \le a_3) \\
&= \Pr(a_1 < X_1 \le b_1, X_2 \le b_2, X_3 \le b_3) - \Pr(a_1 < X_1 \le b_1, X_2 \le a_2, X_3 \le b_3) \\
&\quad - \{\Pr(a_1 < X_1 \le b_1, X_2 \le b_2, X_3 \le a_3) + \Pr(a_1 < X_1 \le b_1, X_2 \le a_2, X_3 \le a_3)\} \\
&= \Pr(X_1 \le b_1, X_2 \le b_2, X_3 \le b_3) - \Pr(X_1 \le a_1, X_2 \le b_2, X_3 \le b_3) - \Pr(X_1 \le b_1, X_2 \le a_2, X_3 \le b_3) \\
&\quad + \Pr(X_1 \le a_1, X_2 \le a_2, X_3 \le b_3) - \Pr(X_1 \le b_1, X_2 \le b_2, X_3 \le a_3) + \Pr(X_1 \le a_1, X_2 \le b_2, X_3 \le a_3) \\
&\quad + \Pr(X_1 \le b_1, X_2 \le a_2, X_3 \le a_3) - \Pr(X_1 \le a_1, X_2 \le a_2, X_3 \le a_3) \\
&= F_{\underline{X}}(b_1, b_2, b_3) - F_{\underline{X}}(a_1, b_2, b_3) - F_{\underline{X}}(b_1, a_2, b_3) + F_{\underline{X}}(a_1, a_2, b_3) - F_{\underline{X}}(b_1, b_2, a_3) + F_{\underline{X}}(a_1, b_2, a_3) \\
&\quad + F_{\underline{X}}(b_1, a_2, a_3) - F_{\underline{X}}(a_1, a_2, a_3) = \sum_{k=0}^{3} (-1)^k \sum_{\underline{z} \in \Delta_{k,3}([\underline{a}, \underline{b}])} F_{\underline{X}}(\underline{z}).
\end{aligned}
$$

The proof can be completed using method of induction. □

The following theorem provides a technique to find marginal distributions.

Theorem 15.14. *Let $F(x_1, x_2, \ldots, x_p)$, $\underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p$ be a d.f. of p-dimensional random vector $\underline{X} = (X_1, X_2, \ldots, X_p)$. Then the marginal distribution function of $\underline{Y} = (X_1, X_2, \ldots, X_{p-1})$ is*

$$G(x_1, x_2, \ldots, x_{p-1}) = \lim_{t \to \infty} F(x_1, x_2, \ldots, x_{p-1}, t), \;\; \underline{y} = (x_1, x_2, \ldots, x_{p-1}) \in \mathbb{R}^{p-1}.$$

Proof. For $\underline{y} = (x_1, x_2, \ldots, x_{p-1}) \in \mathbb{R}^{p-1}$

$$
\begin{aligned}
G(x_1, x_2, \ldots, x_{p-1}) &= \Pr(X_1 \le x_1, \ldots, X_{p-1} \le x_{p-1}) \\
&= \Pr(X_1 \le x_1, \ldots, X_{p-1} \le x_{p-1}, X_p < \infty) \\
&= \Pr\left(\bigcup_{t=1}^{\infty} \{X_1 \le x_1, \ldots, X_{p-1} \le x_{p-1}, X_p \le t\} \right) \\
&= \lim_{t \to \infty} \Pr(X_1 \le x_1, \ldots, X_{p-1} \le x_{p-1}, X_p \le t) = \lim_{t \to \infty} F(x_1, x_2, \ldots, x_{p-1}, t).
\end{aligned}
$$

This completes the proof. □

Theorem 15.15. *Let $\underline{X} = (X_1, X_2, \ldots, X_p)$ be a p-dimensional random vector with d.f. $F(\cdot)$. Then*

(a) $\lim_{\substack{x_i \to \infty \\ i=1,2,\ldots,p}} F(x_1, x_2, \ldots, x_p) = 1,$

(b) for each $i = 1, 2, \ldots, p$, $\lim_{x_i \to -\infty} F(x_1, x_2, \ldots, x_p) = 0,$

(c) $F(\underline{x})$ is right continuous in each argument (keeping other arguments fixed),

(d) for each rectangle $(\underline{a}, \underline{b}] \subseteq \mathbb{R}^p$

$$\sum_{k=0}^{p} (-1)^k \sum_{\underline{z} \in \Delta_{k,p}((\underline{a}, \underline{b}])} F(\underline{z}) \ge 0.$$

Conversely, any function $G : \mathbb{R}^p \to [0, 1]$ satisfying conditions $(a) - (d)$ above is a d.f. of some p-dimensional random vector.

Proof. For simplicity, we provide the proof for $p = 2$.

(a) Note that

$$\begin{aligned}\lim_{x_1\to\infty, x_2\to\infty} F(x_1, x_2) &= \lim_{x_1\to\infty, x_2\to\infty} \Pr(\{X_1 \le x_1, X_2 \le x_2\}) \\ &= \lim_{n\to\infty} \Pr(\{X_1 \le n, X_2 \le n\}), \quad \text{(since limit exists)} \\ &= \Pr(\bigcup_{n=1}^{\infty} \{X_1 \le n, X_2 \le n\}) = \Pr(\{X_1 < \infty, X_2 < \infty\}) = 1.\end{aligned}$$

(b) For fixed $x_2 \in \mathbb{R}$,

$$\begin{aligned}\lim_{x_1\to-\infty} F(x_1, x_2) &= \lim_{n\to\infty} \Pr(\{X_1 \le -n, X_2 \le x_2\}) \\ &= \Pr(\bigcap_{n=1}^{\infty} \{X_1 \le -n, X_2 \le x_2\}) = \Pr(\phi) = 0.\end{aligned}$$

Similarly, $\lim_{x_2\to-\infty} F(x_1, x_2) = 0$.

(c) Let $\{h_n\}_{n\ge1}$ be a sequence in $\mathbb{R}$ such that $h_n \downarrow 0$. Then for $(x_1, x_2) \in \mathbb{R}^2$

$$\begin{aligned}\lim_{n\to\infty} F(x_1 + h_n, x_2) &= \lim_{n\to\infty} \Pr(\{X_1 \le x_1 + h_n, X_2 \le x_2\}) \\ &= \lim_{n\to\infty} \Pr\left(\left\{X_1 \le x_1 + \frac{1}{n}, X_2 \le x_2\right\}\right), \quad \text{(as limit exists)} \\ &= \Pr\left(\bigcap_{n=1}^{\infty} \left\{X_1 \le x_1 + \frac{1}{n}, X_2 \le x_2\right\}\right) = \Pr(\{X_1 \le x_1, X_2 \le x_2\}) = F(x_1, x_2),\end{aligned}$$

i.e. for every fixed $x_2 \in \mathbb{R}$, $F(x_1, x_2)$ is right continuous in $x_1 \in \mathbb{R}$. Similarly, it can be shown that for every fixed $x_1 \in \mathbb{R}$, $F(x_1, x_2)$ is right continuous in $x_2 \in \mathbb{R}$.

(d) For $-\infty < a_1 < b_1 < \infty$ and $-\infty < a_2 < b_2 < \infty$, we have

$$\begin{aligned}\sum_{k=0}^{2} (-1)^k \sum_{\underline{z}\in\Delta_{k,p}((\underline{a},\underline{b}])} F(\underline{z}) &= F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2) \\ &= P(a_1 < X_1 \le b_1,\ a_2 < X_2 \le b_2) \ge 0.\end{aligned}$$

This completes the proof. □

Remark 15.16. *(a) For* $p = 1$*, (d) of the above theorem reduces to* $F(b) - F(a) \ge 0, \forall\ -\infty < a < b < \infty$*, i.e.,* $F(\cdot)$ *is monotone on* $\mathbb{R}$*.*

(b) $F(\cdot)$ *is clearly non-decreasing in each argument.*

15.1.1. Independent Random Variables

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

Definition 15.17. *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 15.18. *(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 15.19. *For any positive integer p (≥ 2) the random variables $X_1, X_2, \ldots, X_p$ are independent iff*

$$F(x_1, x_2, \ldots, x_p) = \prod_{i=1}^{p} F_{X_i}(x_i)\ \forall\ \underline{x} = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p, \tag{15.1}$$

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 (15.1) holds. Conversely suppose that (15.1) 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, $j = 1, 2, \ldots, q$. $\square$