# Multivariate random vectors

$\underset{\sim}{X} = (X_1, \ldots, X_p)'$ $\quad X_i$ is a r.v. on $(\Omega, \mathcal{F}, P)$ say

$\forall \omega \in \Omega$; assign $\underset{\sim}{X}(\omega) = \begin{pmatrix} X_1(\omega) \\ \vdots \\ X_p(\omega) \end{pmatrix}$

## Dist$^n$ f$^n$ or joint dist$^n$ f$^n$

$\underset{\sim}{X}_{p \times 1}$ is a random vector

$$F_{\underset{\sim}{X}}(x_1, \ldots, x_p) = P(\omega : X_1(\omega) \le x_1, X_2(\omega) \le x_2, \ldots X_p(\omega) \le x_p)$$

$$= P(X_1 \le x_1, X_2 \le x_2, \ldots, X_p \le x_p)$$

Note that $F_{\underset{\sim}{X}}(\cdot)$ as defined above satisfies

(i) $\lim\limits_{\min_i x_i \to \infty} F_X(\underset{\sim}{x}) = 1$

(ii) $\lim\limits_{\text{any } x_i \to -\infty} F_{\underset{\sim}{X}}(\underset{\sim}{x}) = 0 \quad$ for $i = 1, \ldots, p$

(iii) $F_{\underset{\sim}{X}}(\underset{\sim}{x})$ is non decreasing in each argument

(iv) $F_X(\underset{\sim}{x})$ is right continuous in each argument

**Remark:** $p = 2$

$$\lim_{x_2 \to \infty} F_{\underset{\sim}{X}}(x_1, x_2) = F_{\underset{\sim}{X}}(x_1, \infty) = F_{X_1}(x_1)$$

marginal dist$^n$ f$^n$ of $X_1$

In general for any $K = 1, \ldots, p$

$$\lim_{x_K \to \infty} F_{\underset{\sim}{X}}(x_1, \ldots \underset{K^{th}}{\uparrow} \ldots x_p) = F_{X_1, \ldots X_{K-1}, X_{K+1}, \ldots, X_p}(x_1, \ldots x_{K-1}, x_{K+1}, \ldots x_p)$$

marginal joint d.f. of $(X_1, \ldots X_{K-1}, X_{K+1}, \ldots X_p)$

Remark :

$P(\underset{\sim}{X} \in B_p)$ can be expressed through joint d.f.

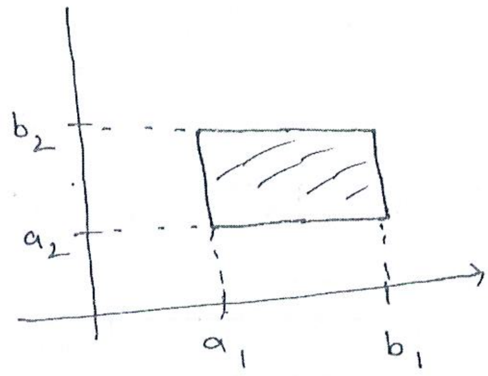
$p$-dimensional semi closed rectangle of the form

$(a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_k, b_k]$

$a_i < b_i$ for $i = 1, \ldots, k$

Consider $p = 2$, to have a feel

$a_1 < b_1$ ; $a_2 < b_2$




$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2)$

$= P(X_1 \leq b_1, X_2 \leq b_2)$

$- P(X_1 \leq a_1, X_2 \leq b_2)$

$- P(X_1 \leq b_1, X_2 \leq a_2)$

$+ P(X_1 \leq a_1, X_2 \leq a_2)$

$$= F_{X_1, X_2}(b_1, b_2) - F_{X_1, X_2}(a_1, b_2) - F_{X_1, X_2}(b_1, a_2) + F_{X_1, X_2}(a_1, a_2)$$

Remark : The four conditions that $F_{\underset{\sim}{X}}(\cdot)$ satisfies (i)-(iv) are not n.s.c. for a $f^n$ to be d.f. of random vector.

We additionally need condition (for $p = 2$) that

$$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) \geq 0 \quad \forall \; \begin{matrix} a_1 < b_1 \\ a_2 < b_2 \end{matrix}$$

# Discrete random vector

**Def$^n$:** A random vector $\underset{\sim}{X} = (X_1, \ldots, X_p)'$ is said to be discrete

If there exist a countable set $E \subset \mathbb{R}^p \ni$ (finite or countably infinite)

$$P(\underset{\sim}{X} \in E) = 1$$

p.m.f. or joint p.m.f of $(X_1, \ldots X_p)$

Let $E = \{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots\}$ ; $\underset{\sim}{e}_1 \in \mathbb{R}^p$

$$p_X(\underline{x}) = P(\underline{X} = \underset{\sim}{x}) = \begin{cases} P(X = \underset{\sim}{e}_i), & \text{If } \underset{\sim}{x} = \underset{\sim}{e}_i, \; i = 1, 2, \ldots \\ 0, & \text{If } \underset{\sim}{x} \notin E \end{cases}$$

i.e. $p_X(\underline{x}) = P(X_1 = x_1, \ldots, X_p = x_p)$.

Consider a bivariate setup $\underline{p = 2}$ case

(i) marginal p.m.f. of $X_i$ can be obtained by summing over possible values of the other variable

e.g $P(X_1 = x_i) = \sum_y P(X_1 = x_i, X_2 = y)$ ; $i = 1, 2, \ldots$

Sly $P(X_2 = y_j) = \sum_x P(X_1 = x, X_2 = y_j)$ ; $j = 1, 2, \ldots$

(ii) Conditional dist$^n$ of $X_1$ given $X_2$ or $X_2$ given $X_1$

conditional p.m.f of $X_2$ given $X_1$

$$p_{X_2 | X_1 = x_i}(y | x_i) = \frac{P(X_2 = y, X_1 = x_i)}{P(X_1 = x_i)}$$

For each level of fixed $x_i$, we get a conditional dist$^n$

Sly $$p_{X_1 | X_2 = y_j}(x | y_j) = \frac{P(X_1 = x, X_2 = y_j)}{P(X_2 = y_j)}$$

For a $p$-dimensional random vector $\underset{\sim}{X} = (X_1, \cdots X_p)$, we can obtain $p$ ~~~~ 1-variate marginals for $X_1, X_2, \cdots X_p$

$\binom{p}{2}$ 2-variate joint marginals for $(X_i, X_j)$

and so on

One can obtain conditional p.m.f. of $(X_{i_1}, X_{i_2}, \cdots, X_{i_r})$ given the remaining variables ~~~~ or any ~~~~ subset of the remaining variables.

<u>Independence</u>: Discrete random variables ~~~~ $X_1, \cdots X_p$ are independent iff the joint p.m.f can be expressed as

$$P(X_1 = x_1, \cdots, X_p = x_p) = \prod_{i=1}^{p} P(X_i = x_i)$$

$$\forall (x_1, \cdots x_p) \in E$$

Note that in such a case conditional p.m.f s will be identical to unconditional p.m.f.s.

<u>Example of a multivariate discrete dist$^n$</u>

Consider a random experiment with 3 mutually exclusive and exhaustive outcomes, $A_1, A_2$ & $A_3$ with probabilities $\theta_1, \theta_2, \theta_3$, respectively. Repeat the trials $n$ times

Define

$X_1$: number of times $A_1$ occurs out of $n$ trials

$X_2$: - - - - - $A_2$ occurs out of $n$ trials

$X_3$: - - - - - $A_3$ occurs out of $n$ trials let $(x_1, x_2, x_3)$ denote the observed count in $n$ trials

$$x_i \geq 0, \quad x_i \leq n \quad \forall i = 1, 2, 3 \quad \& \quad \sum_{1}^{3} x_i = n$$

$$E = \left\{ (x_1, x_2, x_3) : 0 \leq x_i \leq n, \ \sum_{i=1}^{3} x_i = n \right\}$$ - finite number of points

jt p.m.f.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{n!}{x_1! \, x_2! \, x_3!} \theta_1^{x_1} \theta_2^{x_2} \theta_3^{x_3}; \quad 0 \leq x_i \leq n, \ \sum x_i = n$$

Note that $x_3 = n - x_1 - x_2$ and

$$\theta_3 = 1 - \theta_1 - \theta_2$$

$$P(X_1 = x_1, X_2 = x_2) = \frac{n!}{x_1! \, x_2! \, (n - x_1 - x_2)!} \theta_1^{x_1} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{n - x_1 - x_2}$$

$$x_1, x_2 \geq 0$$
$$x_1 + x_2 \leq n$$

$$= 0, \quad \text{o/w}$$

$(X_1, X_2)$ is said to follow a trinomial dist$^n$ $(n, \theta_1, \theta_2)$

Marginal dist$^n$s:

Marginal p.m.f. of $X_1$:

$$P(X_1 = x_1) = \sum_{x_2 = 0}^{n - x_1} \frac{n!}{x_1! \, (n - x_1)!} \theta_1^{x_1} \frac{(n - x_1)!}{x_2! \, (n - x_1 - x_2)!} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{n - x_1 - x_2}$$

$$= \binom{n}{x_1} \theta_1^{x_1} (1 - \theta_1 - \theta_2 + \theta_2)^{n - x_1}$$

$$= \binom{n}{x_1} \theta_1^{x_1} (1 - \theta_1)^{n - x_1}.$$

i.e. $X_1 \sim Bin(n, \theta_1)$

Sly $X_2 \sim Bin(n, \theta_2)$