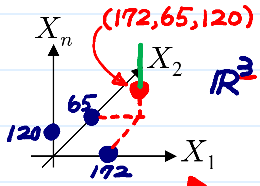
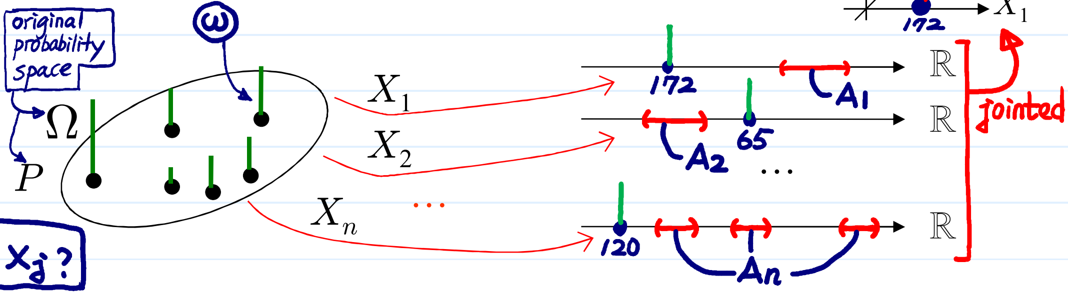
Jointly Distributed Random Variables

• Recall. In Chapters 4 and 5, focus on *univariate* random variable.

➢However, often a single experiment will have more than one random variables which are of interest.



e.g. ω: a sample of k persons
X_1: average height
X_2: average weight
X_3: average blood pressure
How X_i {related to / associated with} X_j ?



➢Definition. Given a sample space Ω and a probability measure P defined on the subsets of Ω, random variables

$$X_1, X_2, \ldots, X_n: \Omega \to \mathbb{R}$$

are said to be *jointly distributed.*

must be maps defined on "same" sample space Ω

Ω is critical in discussing issues like:
(1) $Y(\omega)=g(X(\omega))$
$Y(\omega)>X(\omega)$
(2) $\{X_n<r\}=\{Y_r>n\}$
(3) count process $Y_t(\omega)$
(4) …

univariate $\mathbb{R}^1$ ↕ joint $\mathbb{R}^n$

▪ We can regard n jointly distributed r.v.'s as a *random vector*

cf random variable

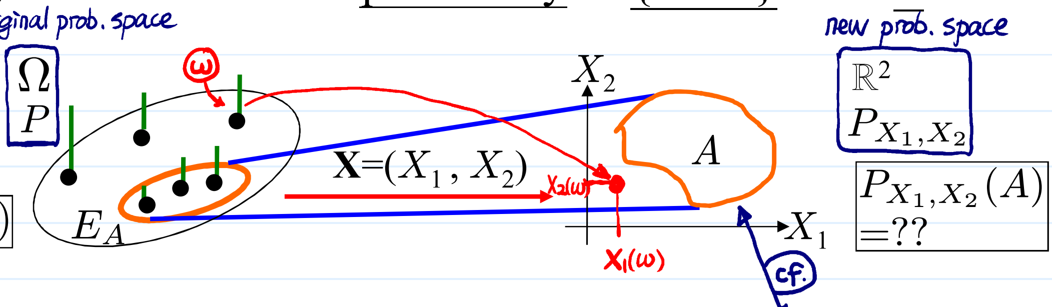
$$\mathbf{X}=(X_1(\omega), \ldots, X_n(\omega)): \Omega \to \mathbb{R}^n.$$

⦿**Q**: For $A\subset\mathbb{R}^n$, how to define the probability of $\{\mathbf{X}\in A\}$ from P?

cf. the graph in LNp.5-4

A occurs ⇔ E_A occurs

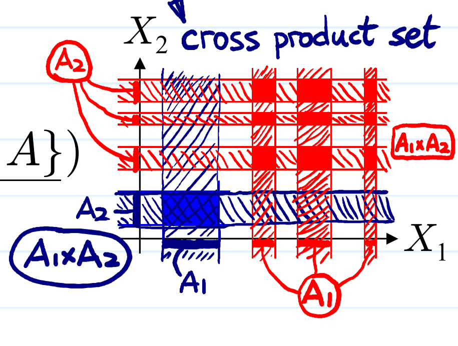
$\boxed{P_{X_1,X_2}(A) = P(E_A)}$



➢For $A\subset\mathbb{R}^n$,

$$P_{X_1,\ldots,X_n}(A) = P(\underbrace{\{\omega\in\Omega \mid (X_1(\omega),\ldots,X_n(\omega))\in A\}}_{E_A})$$

original prob. measure



➢For $A_i\subset\mathbb{R}$, $i=1, \ldots, n$,

$$P_{X_1,\ldots,X_n}(X_1\in A_1, \cdots, X_n\in A_n)$$
$$= P(\{\omega\in\Omega \mid X_1(\omega)\in A_1\}\cap\cdots\cap\{\omega\in\Omega \mid X_n(\omega)\in A_n\})$$

(the commas mean "and")

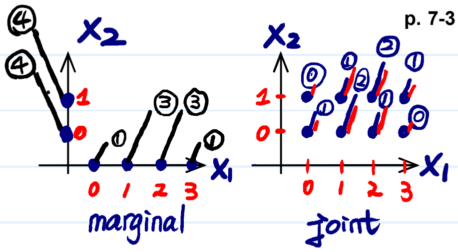
➢Definition. The probability measure of $\mathbf{X}$ ($P_{\mathbf{X}}$, defined on subsets of $\mathbb{R}^n$) is called the *joint distribution* of $X_1, \ldots, X_n$. The probability measure of X_i (P_{X_i}, defined on subsets of $\mathbb{R}$) is called the *marginal distribution* of X_i.

聯合分配

邊際分配

$i=1,2,\cdots,n.$

- **Q**: Why need joint distribution? Why are marginal distributions not enough?



➢ Example (Coin Tossing, Toss a fair coin 3 times, LNp.5-3). $P(\{h,t,t\}\cup\{t,h,t\}\cup\{t,t,h\})=3/8$

$P(\{t,t,t\}\cup\{t,h,t\}\cup\{t,t,h\}\cup\{t,h,h\})=4/8=1/2$

X_2: # of head on 1st toss	X_1: total # of heads			
	0 (1/8)	**1** (3/8)	**2** (3/8)	**3** (1/8)
0 (1/2)	1/8 [1/16]	2/8 [3/16]	1/8 [3/16]	0 [1/16]
1 (1/2)	0 [1/16]	1/8 [3/16]	2/8 [3/16]	1/8 [1/16]

$P(\{h,h,h\})=1/8$

- blue numbers: joint distribution of X_1 and X_2
- (black numbers): marginal distributions
- [red numbers]: joint distribution of another (X_1', X_2'): $\Omega' \to \mathbb{R}^2$
- Some findings: $\Omega=\{(h,h,h),(h,h,t),(h,t,h),(t,h,h),(h,t,t),(t,h,t),(t,t,h),(t,t,t)\}$, each $1/8$
  - When joint distribution is given, its corresponding marginal distributions are known, e.g.,

by the law of total probability

  - ◆ $P(X_1=i)=P(X_1=i, X_2=0)+P(X_1=i, X_2=1)$, $i=0, 1, 2, 3$.

  - (X_1, X_2) and (X_1', X_2') have identical marginal distributions but different joint distributions.
    - ◆ When the marginal distributions are given, the corresponding joint distribution is still unknown. There could be many possible different joint distributions. (A special case: $X_1, \ldots, X_n$ are independent.)

When $X_1, \cdots, X_n$ are independent, their joint dist. can be obtained from their marginal dist. (future lecture)

  - Joint distribution offers more information, e.g.,
    - ◆ When not observing X_1, the distribution of X_2 is: $P(X_2=0)=\boxed{1/2}$, $P(X_2=1)=\boxed{1/2}$ ⇒ marginal distribution
    - ◆ When X_1 was observed, say $X_1=1$, the distribution of X_2 is: $P(X_2=0|X_1=1)=(2/8)/(3/8)=\boxed{2/3}$ and $P(X_2=1|X_1=1)=(1/8)/(3/8)=\boxed{1/3}$ ⇒ the calculation requires the knowing of joint distribution

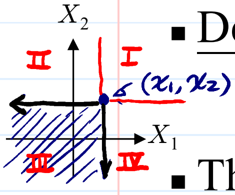
$\dfrac{P(\{\omega|X_2(\omega)=0\}\cap\{\omega|X_1(\omega)=1\})}{P(\{\omega|X_1(\omega)=1\})}$

maps to $\mathbb{R}^n$

c.f. univariate case

- We can characterize the $P_{\mathbf{X}}$ joint distribution of $\mathbf{X}$ in terms of its
  1. Joint Cumulative Distribution Function (joint cdf)
  2. Joint Probability Mass (Density) Function (joint pmf or pdf)
  3. Joint Moment Generating Function (joint mgf, Chapter 7)

➢ Joint Cumulative Distribution Function $F_{\mathbf{X}}: \mathbb{R}^n \to \mathbb{R}^1$



- **Definition.** The joint cdf of $\mathbf{X}=(X_1, \ldots, X_n)$ is defined as

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = P(X_1 \le x_1, \underset{\text{and}(\cap)}{} X_2 \le x_2, \underset{\text{and}(\cap)}{} \ldots, \underset{\text{and}(\cap)}{} X_n \le x_n).$$

- **Theorem.** Suppose that $F_{\mathbf{X}}$ is a joint cdf. Then,

(1) in LNp.5-9 — cf. (i) $0 \le F_{\mathbf{X}}(x_1, \ldots, x_n) \le 1$, for $-\infty < x_i < \infty$, $i=1, \ldots, n$.

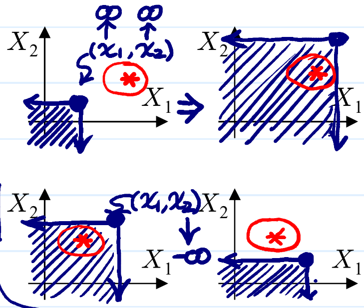
cf. (ii) $\lim_{x_1, x_2, \cdots, x_n \to \infty} F_{\mathbf{X}}(x_1, \ldots, x_n) = 1$

(4) in LNp.5-10: $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$

Proof. Let $z_{im} \uparrow \infty$ ($m \to \infty$), $1 \le i \le n$. ← $\mathbf{z}_m = (z_{1m}, \cdots, z_{nm})$

Let $A_m = (-\infty, z_{1m}] \times \cdots \times (-\infty, z_{nm}]$.

Then, $A_m \uparrow \mathbb{R}^n \Rightarrow \lim \underbrace{P(A_m)}_{F_{\mathbf{X}}(z_{1m}, \cdots, z_{nm})} = P(\mathbb{R}^n) = 1$. (by monotone convergence Thm (LNp.3-17))



cf. (iii) For any $i \in \{1, \ldots, n\}$,

$$\lim_{x_i \to -\infty} F_{\mathbf{X}}(x_1, \ldots, x_n) = 0.$$

Proof. Let $z_{im} \downarrow -\infty$ ($m \to \infty$), for some i.

Let $A_m = (-\infty, \underset{\text{fixed}}{x_1}] \times \cdots \times (-\infty, \underset{\downarrow -\infty}{z_{im}}] \times \cdots \times (-\infty, \underset{\text{fixed}}{x_n}]$

Then, $A_m \downarrow \emptyset \Rightarrow \lim \overbrace{P(A_m)}^{=F_{\mathbf{X}}(x_1, \cdots, z_{im}, \cdots, x_n)} = P(\emptyset) = 0$.