◆ Marginal Distribution. Suppose that

$(X_1, \ldots, X_m) \sim$ Multinomial$(n, m, p_1, \ldots, p_k, p_{k+1}, \ldots, p_m)$.

For $1 \le k < m$, the distribution of

$$(X_1, \ldots, X_k, \underbrace{X_{k+1} + \cdots + X_m}_{= n-(X_1+\cdots+X_k)})$$

is Multinomial$(n, k+1, p_1, \ldots, p_k, p_{k+1}+\cdots+p_m)$.

In particular, $X_i \sim$ Binomial(n, p_i).

$p_1, p_2, \ldots$ m types $p_{k+1}, \ldots, p_m$: $1, 2, \ldots, k, k+1, \ldots, m$ with $X_1, X_2, \ldots, X_k$; $k+1, \ldots, m$ → a new type $(k+1)$: $p_{k+1}+\cdots+p_m$, $X_{k+1}+\cdots+X_m$. can be any k r.v.'s. For its proof, use (e) in LNp.7-8 & (∗) in LNp.7-16 (exercise). $1, 2, \ldots, m$: X_1, a new type, $X_2+\cdots+X_m$; $(X_1, X_2+\cdots+X_m) \sim$ multinomial$(n, \underline{\underline{2}}, p_1, p_2+\cdots+p_m)$.

$\sum \binom{n-x_1-\cdots-x_k}{x_{k+1}, \ldots, x_m} \times p_{k+1}^{x_{k+1}} \cdots p_m^{x_m} = (p_{k+1}+\cdots+p_m)^{n-x_1-\cdots-x_k}$

◆ Mean and Variance.

$$E(X_i) = np_i \text{ and } Var(X_i) = np_i(1-p_i)$$

for $i = 1, \ldots, m$.

➢ Example. examine whether it's a joint pdf (Ex)

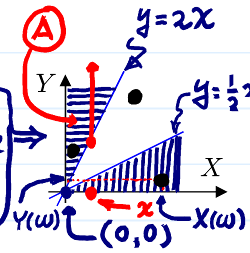
▪ Suppose that the joint pdf of 2 continuous r.v.'s (X, Y) is

$$f(x, y) = \begin{cases} \lambda^2 e^{-\lambda(x+y)}, & x \ge 0,\ y \ge 0, \\ 0, & \text{otherwise.} \end{cases}$$

Note: $f(x,y) = \underbrace{(\lambda e^{-\lambda x})}_{f_1(x)}\underbrace{(\lambda e^{-\lambda y})}_{f_2(y)}$

Q: $P(Y(\omega) \ge 2X(\omega)$ or $X \ge 2Y) = ??$

▪ The event $\{Y \ge 2X\} \cup \{X \ge 2Y\}$ is



mutually exclusive except (0,0)

▪ So, $P(Y \ge 2X$ or $X \ge 2Y) = P(Y \ge 2X) + P(X \ge 2Y) = 2/3$ because

∵ mutually exclusive except (0,0) & $P((0,0)) = 0$

$P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$

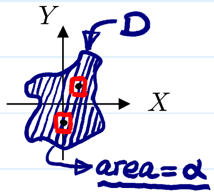
$$P(Y \ge 2X) = \int_0^\infty \left[\int_{2x}^\infty \lambda^2 e^{-\lambda(x+y)}\, dy \right] dx$$

$$= \int_0^\infty -\lambda e^{-\lambda(x+y)} \Big|_{y=2x}^{\infty} dx = \int_0^\infty \lambda e^{-3\lambda x}\, dx$$

$$= (-1/3)e^{-3\lambda x} \Big|_{x=0}^{\infty} = 1/3.$$

and similarly, we can get $P(X \ge 2Y) = 1/3$ (exercise).

➢ Example. Consider two continuous r.v.'s X and Y.



▪ Uniform Distribution over a region D. If $D \subset \mathbb{R}^2$ and $0 < \alpha = \text{Area}(D) < \infty$, then

cf. Indicator function in LNp.5-13

$$f(x, y) = c \cdot \mathbf{1}_D(x, y)$$

indicator function, $\mathbb{1}_D(x,y) = \begin{cases} 1, & \text{if } (x,y) \in D \\ 0, & \text{if } (x,y) \notin D \end{cases}$

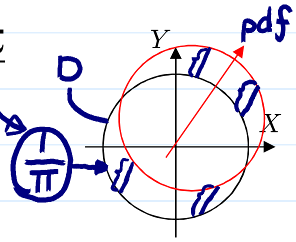
is a joint pdf when $c = 1/\alpha$, called the uniform pdf over D.

$\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = c \iint_{\mathbb{R}^2} \mathbb{1}_D(x,y)\,dx\,dy = c \cdot \iint_D 1\,dx\,dy = c \cdot \text{area}(D) = 1$

▪ Let $D = \{(x, y): x^2+y^2 \le 1\}$, then $\alpha = \text{Area}(D) = \pi$ and

$$f(x, y) = \frac{1}{\pi} \mathbf{1}_D(x, y)$$

is a joint pdf.



- Marginal distribution. The marginal pdf of X is

still uniform?

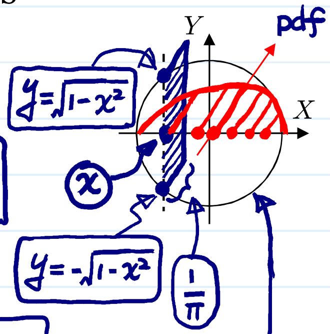
Thm in LNp.7-10

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy = \frac{2}{\pi}\sqrt{1-x^2}$$

X is not uniform

for $-1 \le x \le 1$, and $f_X(x)=0$, otherwise.

(exercise: Find the marginal distribution of Y.)



not a cross product set (check LNp.7-2)

❖ **Reading**: textbook, Sec 6.1

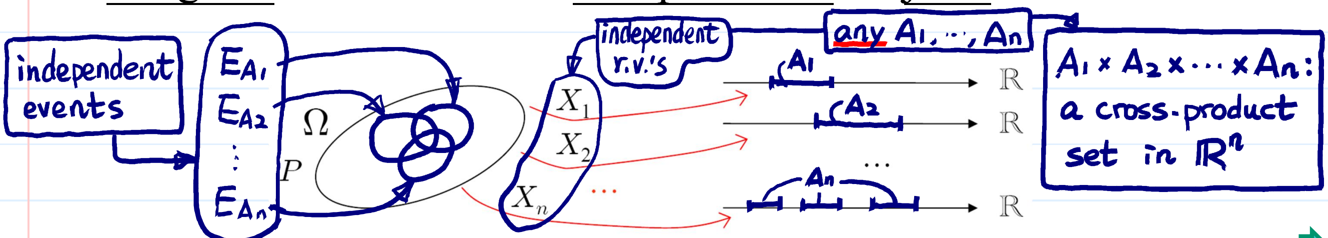
Recall. sum of indep. Bernoullis → binomial, ⋯

Independent Random Variables

c.f. independent events (LNp.4-19~30)

- Recall.
  - ➢ If joint distribution is given, marginal distributions are known.
  - ➢ The converse statement does not hold in general.
  - ➢ However, when random variables are independent,

    marginal distributions + independence ⇒ joint distribution.



- **Definition.** The n jointly distributed r.v.'s $X_1, \ldots, X_n$ are called (*mutually*) *independent* if and only if for $\boxed{any}$ (measurable) sets $A_i \subset \mathbb{R}$, $i=1, \ldots, n$, the events

for calculation purpose

$A_1 \times A_2 \times \cdots \times A_n$: a cross product set in $\mathbb{R}^n$

$A_1, \cdots, A_n \subset \mathbb{R}^1$ ∩

$$\underbrace{\{X_1 \in A_1\}}_{E_{A_1}}, \ldots, \underbrace{\{X_n \in A_n\}}_{E_{A_n}}$$

$E_{A_1}, \cdots, E_{A_n} \subset \Omega$

are (mutually) independent. That is,

c.f. joint dist.

$E_{A_{i_1}}$ and $E_{A_{i_2}}$ and and $E_{A_{i_k}}$

a cross product set in $\mathbb{R}^k$ (LNp.7-2)

marginal dist.

can define prob. of region in (vi) (LNp.7-6) when they're independent

$$P(X_{i_1} \in A_{i_1}, X_{i_2} \in A_{i_2}, \cdots, X_{i_k} \in A_{i_k})$$
$$= P(X_{i_1} \in A_{i_1}) \times P(X_{i_2} \in A_{i_2}) \times \cdots \times P(X_{i_k} \in A_{i_k}),$$

for *any* $1 \le i_1 < i_2 < \cdots < i_k \le n$; $k=2, \ldots, n$.

Actually, it's enough to examine $k=n$. For $k<n$, let $A_j = \mathbb{R}^1$, $j \ne i_1, \cdots, i_k$. (*)

- If $X_1, \ldots, X_n$ are independent, for $1 \le k < n$, (any k)

for interpretation purpose

$$P(X_{k+1} \in A_{k+1}, \ldots, X_n \in A_n \mid X_1 \in A_1, \ldots, X_k \in A_k)$$
$$= P(X_{k+1} \in A_{k+1}, \ldots, X_n \in A_n)$$

exercise $\gtrless$

could be
- a point in $\mathbb{R}^1$
- some points in $\mathbb{R}^1$
- a region in $\mathbb{R}^1$
- a larger region in $\mathbb{R}^1$
- ⋯

provided that $P(X_1 \in A_1, \ldots, X_k \in A_k) > 0$.

- In other words, the values of $X_1, \ldots, X_k$ do *not* carry any information about the distribution of $X_{k+1}, \ldots, X_n$.

- **Theorem** (Factorization Theorem). The random variables $\mathbf{X}=(X_1, \ldots, X_n)$ are independent if and only if one of the following conditions holds. marginal ⇒ joint

(1) $F_{\mathbf{X}}(x_1, \ldots, x_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n)$, where $F_{\mathbf{X}}$ is the joint cdf of $\mathbf{X}$ and F_{X_i} is the marginal cdf of X_i for $i=1,\ldots,n$.

(2) Suppose that $X_1, \ldots, X_n$ are discrete random variables. $p_{\mathbf{X}}(x_1, \ldots, x_n) = p_{X_1}(x_1) \times \cdots \times p_{X_n}(x_n)$, where $p_{\mathbf{X}}$ is the joint pmf of $\mathbf{X}$ and p_{X_i} is the marginal pmf of X_i for $i=1,\ldots,n$.

(3) Suppose that $X_1, \ldots, X_n$ are continuous random variables. $f_{\mathbf{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n)$, where $f_{\mathbf{X}}$ is the joint pdf of $\mathbf{X}$ and f_{X_i} is the marginal pdf of X_i for $i=1,\ldots,n$.

Proof.

independent $\Rightarrow$ (1).

$$\begin{aligned} F_{\mathbf{X}}(x_1, \ldots, x_n) &= P(X_1 \le x_1, \ldots, X_n \le x_n) \\ &= P(X_1 \in (-\infty, x_1], \ldots, X_n \in (-\infty, x_n]) \\ &= P(X_1 \in (-\infty, x_1]) \times \cdots \times P(X_n \in (-\infty, x_n]) \\ &= F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n) \end{aligned}$$

a cross product set in $\mathbb{R}^n$

by the definition of independence

use the property of σ-field

independent $\Leftarrow$ (1). Out of the scope of this couse so skip.

independent $\Rightarrow$ (2).

$$\begin{aligned} p_{\mathbf{X}}(x_1, \ldots, x_n) &= P(X_1 = x_1, \ldots, X_n = x_n) \\ &= P(X_1 \in \{x_1\}, \ldots, X_n \in \{x_n\}) \\ &= P(X_1 \in \{x_1\}) \times \cdots \times P(X_n \in \{x_n\}) \\ &= p_{X_1}(x_1) \times \cdots \times p_{X_n}(x_n) \end{aligned}$$

a cross product set in $\mathbb{R}^n$

by the definition of independence

indep. ⇕

(2) $\Rightarrow$ (1).

by the Thm in LNp.7-9

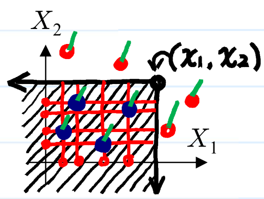
$$F_{\mathbf{X}}(x_1, \ldots, x_n) = \sum_{\substack{(t_1,\ldots,t_n)\in\mathcal{X} \\ t_1 \le x_1, \ldots, t_n \le x_n}} p_{\mathbf{X}}(t_1, \ldots, t_n)$$

$\underset{\sim}{t} = (t_1,\ldots,t_n)$, $\underset{\sim}{t} \in (-\infty, x_1] \times \cdots \times (-\infty, x_n]$ — a cross product set

by (2)

$$= \sum_{\substack{(t_1,\ldots,t_n)\in\mathcal{X} \\ t_1 \le x_1}} \cdots \boxed{\sum_{\substack{(t_1,\ldots,t_n)\in\mathcal{X} \\ t_n \le x_n}}} p_{X_1}(t_1) \times \cdots \times p_{X_n}(t_n)$$

$$= \sum_{\substack{(t_1,\ldots,t_n)\in\mathcal{X} \\ t_1 \le x_1}} p_{X_1}(t_1) \times \cdots \times \sum_{\substack{(t_1,\ldots,t_n)\in\mathcal{X} \\ t_n \le x_n}} p_{X_n}(t_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n)$$



(3) $\Rightarrow$ (1). Thm in LNp.7-11

integration over a cross product set

$$F_{\mathbf{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{\mathbf{X}}(t_1, \ldots, t_n)\, dt_1 \cdots dt_n$$

by (3)

$$= \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{X_1}(t_1) \times \cdots \times f_{X_n}(t_n)\, dt_1 \cdots dt_n$$

$$= \int_{-\infty}^{x_1} f_{X_1}(t_1)\, dt_1 \times \cdots \times \int_{-\infty}^{x_n} f_{X_n}(t_n)\, dt_n = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n)$$

integration region $(-\infty, x_1]$ is irrelevant to $t_2, \cdots, t_n$ cf. graph in LNp.7-17

(3) $\Leftarrow$ (1). Thm in LNp.7-11

$$f_{\mathbf{X}}(x_1,\ldots,x_n) = \frac{\partial^n}{\partial x_1\cdots\partial x_n}F_{\mathbf{X}}(x_1,\ldots,x_n).$$

(check (*) in LNp.7-20)

$F_{X_1,\ldots,X_k}(x_1,\cdots,x_k)$
$= F_{\mathbf{X}}(x_1,\cdots,x_k,\underset{\infty}{x_{k+1}},\cdots,\underset{\infty}{x_n})$
$= F_{X_1}(x_1)\cdots F_{X_k}(x_k) \times \underbrace{F_{X_{k+1}}(x_{k+1})}_{=1}\cdots\underbrace{F_{X_n}(x_n)}_{=1}$

by (1)
$$= \frac{\partial^n}{\partial x_1\cdots\partial x_n}F_{X_1}(x_1)\times\cdots\times F_{X_n}(x_n)$$
$$= \frac{\partial}{\partial x_1}F_{X_1}(x_1)\times\cdots\times\frac{\partial}{\partial x_n}F_{X_n}(x_n) = f_{X_1}(x_1)\times\cdots\times f_{X_n}(x_n)$$

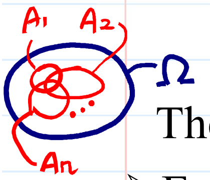
➤Remark. It follows from the Multiplication Law (LNp.4-11) that

$$\begin{aligned}
F_{\mathbf{X}}(x_1,\ldots,x_n) &= P(X_1\le x_1,\ldots,X_n\le x_n)\\
&= P(X_1\le x_1) && (=F_{X_1}(x_1))\\
&\quad\times P(X_2\le x_2|X_1\le x_1) && \left(\overset{?}{=}P(X_2\le x_2)=F_{X_2}(x_2)\right)\\
&\quad\times P(X_3\le x_3|X_1\le x_1,X_2\le x_2) && \left(\overset{?}{=}P(X_3\le x_3)=F_{X_3}(x_3)\right)\\
&\quad\times\cdots\\
&\quad\times P(X_n\le x_n|X_1\le x_1,\ldots,X_{n-1}\le x_{n-1}) && \left(\overset{?}{=}P(X_n\le x_n)=F_{X_n}(x_n)\right)
\end{aligned}$$

X_2 indep of X_1

X_3 indep of X_1, X_2

X_n indep. of $X_1,\cdots,X_{n-1}$



The independence can be established *sequentially*.

➤Example. If $A_1, \ldots, A_n \subset \Omega$ are independent events, then $\mathbf{1}_{A_1},\ldots,\mathbf{1}_{A_n}$, are independent random variables. For example,

indicator functions ~ Bernoulli ($P(A_i)$) discrete

joint pmf $\rightarrow P(\mathbf{1}_{A_1}\overset{0\text{ or }1}{=}1,\mathbf{1}_{A_2}\overset{0\text{ or }1}{=}0,\mathbf{1}_{A_3}\overset{0\text{ or }1}{=}1)$

$$= P(A_1\cap A_2^c\cap A_3) = P(A_1)P(A_2^c)P(A_3)$$

by Thm in LNp.4-24

$$= P(\mathbf{1}_{A_1}=1)P(\mathbf{1}_{A_2}=0)P(\mathbf{1}_{A_3}=1).$$

marginal pmf

product of marginal pmfs

11/14

➤Theorem. If $\mathbf{X}=(X_1, \ldots, X_n)$ are independent and

$$Y_i = g_i(X_i),\ i=1,\ldots,n,$$

then $g_i:\mathbb{R}^1\rightarrow\mathbb{R}^1$

$Y_1, \ldots, Y_n$ are independent.

$\Omega \xrightarrow{X} \mathbb{R}^n \xrightarrow{Y_i=g_i(X_i)} \mathbb{R}^n$

X_1: 生日 $\xleftrightarrow{\text{indep.}}$ X_2: 身高

$X_1 \xrightarrow{g_1} Y_1$: 星座, $X_2 \xrightarrow{g_2} Y_2$: 可否乘車免費

indep

indep.

generalization
$1 = i_0 < i_1 < \cdots < i_k = n$
$Y_1 = g_1(X_1,\ldots,X_{i_1}),$
$Y_2 = g_2(X_{i_1+1},\ldots,X_{i_2}),$
$\cdots$
$Y_k = g_k(X_{i_{k-1}+1},\ldots,X_{i_k}).$

X_1 X_n

different independent r.v.'s

Proof.

Let $A_i(y) = \{x : g_i(x)\le y\}$, $i=1,\ldots,n$, then

$$\begin{aligned}
F_{\mathbf{Y}}(y_1,\ldots,y_n) &= P(Y_1\le y_1,\ldots,Y_n\le y_n)\\
&= P(X_1\in A_1(y_1),\ldots,X_n\in A_n(y_n))\\
&= P(X_1\in A_1(y_1))\times\cdots\times P(X_n\in A_n(y_n))\\
&= P(Y_1\le y_1)\times\cdots\times P(Y_n\le y_n)\\
&= F_{Y_1}(y_1)\times\cdots\times F_{Y_n}(y_n).
\end{aligned}$$

joint cdf

same subset of Ω

a cross product set

$\because X_1,\cdots,X_n$ indep.

same subset of Ω

product of marginal cdfs