Then $f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) \times |J|$, ⟵c.f.⟶ Thm in LNp.6-10~11

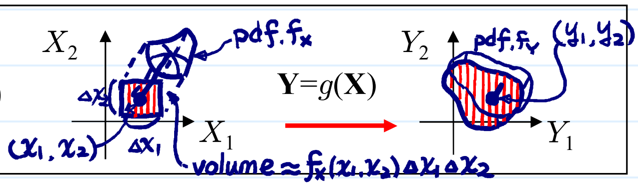
absolute value of J

for $\mathbf{y}$ s.t. $\mathbf{y}=g(\mathbf{x})$ for some $\mathbf{x}$, and $f_{\mathbf{Y}}(\mathbf{y})=0$, otherwise.

12/12

Recall. The question in LNp.6-11

(**Q**: What is the role of $|J|$?)



Proof. $F_{\mathbf{Y}}(y_1, \ldots, y_n) = \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_n} f_{\mathbf{Y}}(t_1, \ldots, t_n)\, dt_n \cdots dt_1$

method of cdf check 1~3 in LNp.7-31

$$= \int \cdots \int_{\substack{(x_1,\ldots,x_n):\\ Y_1 = g_1(x_1,\ldots,x_n)\le y_1\\ \cdots \\ Y_n = g_n(x_1,\ldots,x_n)\le y_n}} f_{\mathbf{X}}(x_1, \ldots, x_n)\, dx_n \cdots dx_1.$$

It then follows from an exercise in advanced calculus that

$$f_{\mathbf{Y}}(y_1, \ldots, y_n) = \frac{\partial^n}{\partial y_1 \cdots \partial y_n} F_{\mathbf{Y}}(y_1, \ldots, y_n)$$

$$= f_{\mathbf{X}}(\underbrace{w_1(\mathbf{y}), \ldots, w_n(\mathbf{y})}_{g^{-1}(\mathbf{y})}) \times |J|.$$

functions of $\mathbf{X}$

- Remark. When the dimensionality of $\mathbf{Y}$ (denoted by k) is less than n, we can choose another $n-k$ transformations $\mathbf{Z}$ such that

$$(\underbrace{\mathbf{Y}}_{k\text{-dim}}, \underbrace{\mathbf{Z}}_{(n-k)\text{-dim}}) = g(\underbrace{\mathbf{X}}_{n\text{-dim}})$$

$\mathbb{R}^n \to \mathbb{R}^n$, g^{-1} exists, g^{-1}: differentiable

satisfy the assumptions in above theorem.

By integrating out the last $n-k$ arguments in the joint pdf of $(\mathbf{Y}, \mathbf{Z})$, the joint pdf of $\mathbf{Y}$ can be obtained.

⟵ by Thm in LNp.7-10

- **Example.** X_1 and X_2 are random variables with joint pdf $f_{\mathbf{X}}(x_1, x_2)$. Find the distribution of $Y_1 = X_1/(X_1+X_2)$. $\equiv g_1(X_1, X_2)$

X_1, X_2: continuous r.v.'s

proportion when $X_1 \ge 0, X_2 \ge 0$

$g=(g_1, g_2): \mathbb{R}^2 \to \mathbb{R}^2$

- Let $Y_2 = X_1 + X_2$, then

$g_2(X_1, X_2)$ add one more transformation

$$x_1 = y_1 y_2 \equiv w_1(y_1, y_2)$$
$$x_2 = y_2 - y_1 y_2 \equiv w_2(y_1, y_2).$$

g^{-1}

Since $\frac{\partial w_1}{\partial y_1} = y_2$, $\frac{\partial w_1}{\partial y_2} = y_1$, $\frac{\partial w_2}{\partial y_1} = -y_2$, $\frac{\partial w_2}{\partial y_2} = 1 - y_1$,

determinant absolute value

$$J = \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1-y_1 \end{vmatrix} = y_2 - \cancel{y_1 y_2} + \cancel{y_1 y_2} = y_2, \text{ and } |J| = |y_2|.$$

Therefore, $f_{\mathbf{Y}}(y_1, y_2) = f_{\mathbf{X}}(y_1 y_2, y_2 - y_1 y_2)|y_2|$,

and, $f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{\mathbf{Y}}(y_1, y_2)\, dy_2$

$$= \int_{-\infty}^{\infty} f_{\mathbf{X}}(y_1 y_2, y_2 - y_1 y_2)|y_2|\, dy_2.$$

$f_{\mathbf{X}}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$

$$\left(\stackrel{\downarrow}{=} \int_{-\infty}^{\infty} f_{X_1}(y_1 y_2) f_{X_2}(y_2 - y_1 y_2)|y_2|\, dy_2\right.$$

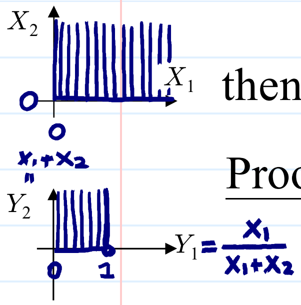
when X_1 and X_2 are independent)

- Theorem. If X_1 and X_2 are independent, and

$$X_1 \sim \text{Gamma}(\alpha_1, \lambda), \quad X_2 \sim \text{Gamma}(\alpha_2, \lambda),$$

then $Y_1 = X_1/(X_1+X_2) \sim \text{Beta}(\alpha_1, \alpha_2)$.

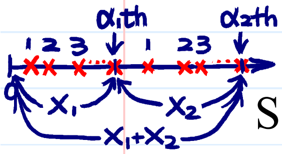
Corollary. $X_1 \sim$ exponential (λ), $X_2 \sim$ exponential (λ) $\Rightarrow Y_1 \sim$ Uniform$(0,1)$



Proof. For $x_1, x_2 \geq 0$, the joint pdf of $\mathbf{X}$ is

$$f_{\mathbf{X}}(x_1, x_2) = \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} x_1^{\alpha_1-1} e^{-\lambda x_1} \times \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} x_2^{\alpha_2-1} e^{-\lambda x_2}$$

$$= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x_1^{\alpha_1-1} x_2^{\alpha_2-1} e^{-\lambda(x_1+x_2)}.$$



So, for $0 \leq y_1 \leq 1$,

$f_{Y_1,Y_2}(y_1,y_2) \propto g_1(y_1) g_2(y_2)$

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{X_1}(y_1 y_2) f_{X_2}(y_2 - y_1 y_2)|y_2|\, dy_2$$

$$= \boxed{\int_0^{\infty}} \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} (y_1 y_2)^{\alpha_1-1}(y_2 - y_1 y_2)^{\alpha_2-1} e^{-\lambda y_2} \cdot y_2\, dy_2$$

$$= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}$$

$$\times \int_0^{\infty} \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)} y_2^{(\alpha_1+\alpha_2)-1} e^{-\lambda y_2}\, dy_2$$

$$= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}$$

pdf of Y_2, i.e. $Y_2 \sim \text{Gamma}(\alpha_1+\alpha_2, \lambda)$

pdf of Gamma$(\alpha_1+\alpha_2, \lambda)$

$E(Y_1) = \frac{\alpha_1}{\alpha_1+\alpha_2}$

pdf of Beta(α_1, α_2)

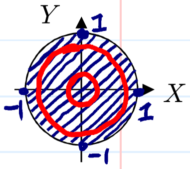
and $f_{Y_1}(y_1) = 0$, otherwise.

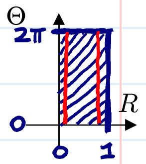
(exercise) Y_1 & Y_2 are independent

- Example. Suppose that X and Y have a uniform distribution over the region $D = \{(x, y): x^2+y^2 \leq 1\}$, i.e., their joint pdf is

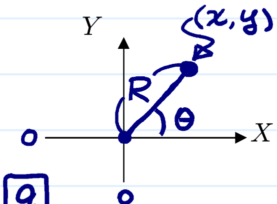
$$f_{X,Y}(x, y) = \frac{1}{\pi} \mathbf{1}_D(x, y).$$

LN p.7-18





Find the joint distribution of (R, Θ) and examine whether R and Θ are independent, where (R, Θ) is the polar coordinate representation of (X, Y), i.e.,



$$X = R\cos(\Theta) \equiv w_1(R, \Theta),$$
$$Y = R\sin(\Theta) \equiv w_2(R, \Theta).$$

g^{-1}

g

$R = \sqrt{x^2+y^2}$
$\Theta = \tan^{-1}\left(\frac{y}{x}\right)$

Since $\frac{\partial w_1}{\partial r} = \cos(\theta)$, $\quad \frac{\partial w_1}{\partial \theta} = -r\sin(\theta)$,
$\frac{\partial w_2}{\partial r} = \sin(\theta)$, $\quad \frac{\partial w_2}{\partial \theta} = r\cos(\theta)$,

$$J = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r\cos^2(\theta) + r\sin^2(\theta) = r,$$

and $|J| = |r| = r$. ← Recall In calculus, for polar transformation, $dxdy = r\,drd\theta$

marginal dist. $\Theta \sim$ Uniform$(0, 2\pi)$
$R \sim$ pdf: $2r\,\mathbf{1}_{(0,1)}(r)$ ← Why?

For $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$, the joint pdf of (R, Θ) is

cross product set

$$f_{R,\Theta}(r, \theta) = f_{X,Y}(r\cos(\theta), r\sin(\theta)) \times |J| = \frac{1}{\pi} r = \frac{1}{2\pi}(2r)$$

and $f_{R,\Theta}(r, \theta) = 0$, otherwise.

- By the theorem in LNp.7-25, (R, Θ) are independent.

- Example. Let $X_1, \ldots, X_n$ be **independent** and **identically distributed** (i.e., **i.i.d.**) exponential(λ). Let

common marginal dist.

$$Y_i = X_1 + \cdots + X_i,\ i = 1, \ldots, n.$$

$g \rightarrow \begin{cases} Y_1 = X_1 \\ Y_2 = X_1 + X_2 \\ \cdots \\ Y_n = X_1 + X_2 + \cdots + X_n \end{cases}$

Find the distribution of $\mathbf{Y} = (Y_1, \ldots, Y_n)$.

marginal (Q: $Y_1, \cdots, Y_n$ indep?)

Note: $Y_1 < Y_2 < \cdots < Y_n$

not a cross product set

[Note. It has been shown that $Y_i \sim$ Gamma(i, λ), i=1, …, n.] LNp. 7-33

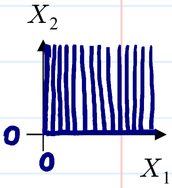
- The joint pdf of $X_1, \ldots, X_n$ is

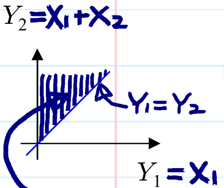
Note: not "x"

$$f_{\mathbf{X}}(x_1, \ldots, x_n) = \prod_{i=1}^n f_{X_i}(x_i)$$
$$= \prod_{i=1}^n \left(\lambda e^{-\lambda x_i}\right) = \lambda^n e^{-\lambda(x_1 + \cdots + x_n)}$$

Note. not "nx"

for $0 \le x_i < \infty$, i=1, …, n.





not a cross product set

$Y_1 < Y_2 < \cdots < Y_n$

g^{-1}

- Since
$$\begin{aligned} x_1 &= y_1 \equiv w_1(y_1, \ldots, y_n), \\ x_2 &= y_2 - y_1 \equiv w_2(y_1, \ldots, y_n), \\ &\cdots \\ x_n &= y_n - y_{n-1} \equiv w_n(y_1, \ldots, y_n), \end{aligned}$$

$\mathbf{Y} = g(\mathbf{X})$
g is one-to-one
$\Rightarrow g^{-1}$ exists

we have
$$\frac{\partial w_i}{\partial y_j} = \begin{cases} 1, & \text{if } j = i, \\ -1, & \text{if } j = i-1, \\ 0, & \text{otherwise}, \end{cases}$$

$$J = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = 1, \text{ and } |J| = 1.$$

- For $0 \le y_1 \le y_2 \le \cdots \le y_{i-1} \le \underset{y}{y_i} \le y_{i+1} \le \cdots \le y_n < \infty$,
$$\begin{aligned} f_{\mathbf{Y}}(y_1, \ldots, y_n) &= f_{\mathbf{X}}(y_1, y_2 - y_1, \ldots, y_n - y_{n-1}) \times |J| \\ &= \lambda^n e^{-\lambda y_n}. \end{aligned}$$

Q: Are $Y_1, \cdots, Y_n$ indep. by the Thm in LNp.7-25?

and $f_{\mathbf{Y}}(y_1, \ldots, y_n) = 0$, otherwise.

check ★ in LNp.7-41

- The marginal pdf of Y_i is

By the Thm in LNp. 7-10

$$\begin{aligned} f_{Y_i}(y) &= \int_0^y \int_{y_1}^y \cdots \int_{y_{i-2}}^y \int_y^\infty \int_{y_{i+1}}^\infty \cdots \int_{y_{n-1}}^\infty \\ &\quad \lambda^n e^{-\lambda y_n}\, dy_n \cdots dy_{i+2}\, dy_{i+1}\, dy_{i-1} \cdots dy_2\, dy_1 \\ &= \int_0^y \int_{y_1}^y \cdots \int_{y_{i-2}}^y \lambda^i e^{-\lambda y}\, dy_{i-1} \cdots dy_2\, dy_1 \\ &= \lambda^i e^{-\lambda y} \frac{y^{i-1}}{(i-1)!}, \end{aligned}$$

a constant

$\Gamma(i)$

pdf of Gamma(i, λ)

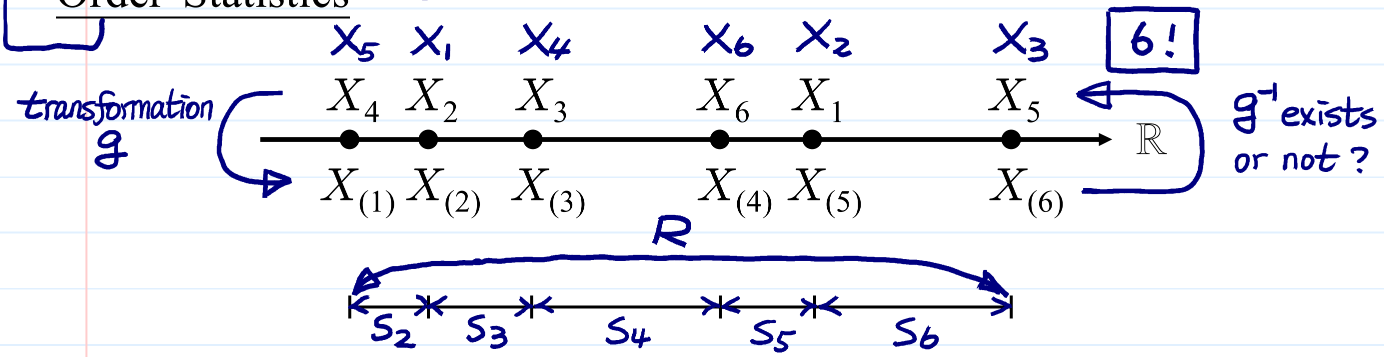
for $y \ge 0$, and $f_{Y_i}(y) = 0$, otherwise.

➢ Method of moment generating function.

- Based on the *uniqueness theorem* of moment generating function to be explained later in Chapter 7
- Especially useful to identify the distribution of sum of independent random variables. ⇒ $X_1, \cdots, X_n$ indep. , $Y = X_1 + \cdots + X_n$
  mgf of $Y = \prod$ mgf of X_i

順序統計量

• Order Statistics → quantile (分位數)



➢ Definition. Let $X_1, \ldots, X_n$ be random variables. We sort the X_i's and denote by

$$X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$$

the *order statistics*. Using the notation,

$X_{(i)}$ = ith-smallest value in $X_1, \ldots, X_n$, i=1, 2, …, n,

$X_{(1)}$ = min($X_1, \ldots, X_n$) is the *minimum*,

$X_{(n)}$ = max($X_1, \ldots, X_n$) is the *maximum*,

e.g. $X_1, \cdots, X_n \sim$ Uniform(a, b)
a, b : unknown
a $X_{(1)}$ b $X_{(n)}$

transformations of order stat. $X_{(1)}, \cdots, X_{(n)}$

$R \equiv X_{(n)} - X_{(1)}$ is called *range*,

$S_j \equiv X_{(j)} - X_{(j-1)}$, j=2, …, n, are called *j*th *spacing*.

Q: What are the joint distributions of various order statistics and their marginal distributions?

12/14

➢ Definition. $X_1, \ldots, X_n$ are called *i.i.d.* (**i**ndependent, **i**dentically **d**istributed) with cdf F/pdf f/pmf p if the random variables $X_1, \ldots, X_n$ are independent and have a common marginal distribution with cdf F/pdf f/pmf p.

Intuition: (1) $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$
(2) $X_{(1)} = x$, $X_{(2)} \ge x \Rightarrow$ conditional dist of $X_{(2)} | X_{(1)} = x$ change with x

- Remark. In the discussion about order statistics, we only consider the case that $X_1, \ldots, X_n$ are i.i.d.

exception
$X_1 \in [0, 1]$
$X_2 \in [2, 3]$
$X_{(1)}, X_{(2)}$ indep

◘ Note. Although $X_1, \ldots, X_n$ are independent, their order statistics $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ are *not independent* in general.