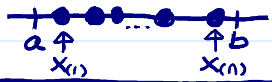
$X_{(i)}$ = ith-smallest value in $X_1, \ldots, X_n$, i=1, 2, …, n,

$X_{(1)}$ = 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



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 *jth spacing*.

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

11/21

➢ 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.

➢ Theorem. Suppose that $X_1, \ldots, X_n$ are i.i.d. with cdf F.

Q: Which of the methods 1, 2, 3 in LNp.7-28 will you choose to identify the dist. of $X_{(1)}$ & $X_{(n)}$?

1. The cdf of $X_{(1)}$ is $1-[1-F(x)]^n$, and the cdf of $X_{(n)}$ is $[F(x)]^n$.

2. If **X** are continuous and F has a pdf f, then the pdf of $X_{(1)}$ is $nf(x)[1-F(x)]^{n-1}$, and the pdf of $X_{(n)}$ is $nf(x)[F(x)]^{n-1}$.

Proof. By the method of cumulative distribution function,

$$1 - F_{X_{(1)}}(x) = P(X_{(1)} > x) = P(X_1 > x, \ldots, X_n > x)$$
$$= P(X_1 > x) \cdots P(X_n > x) = [1 - F(x)]^n.$$

Note not "x_i" (c.f., the notes in LNp.7-41)

∵ independent

$$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P(X_1 \le x, \ldots, X_n \le x)$$
$$= P(X_1 \le x) \cdots P(X_n \le x) = [F(x)]^n.$$

∵ identical

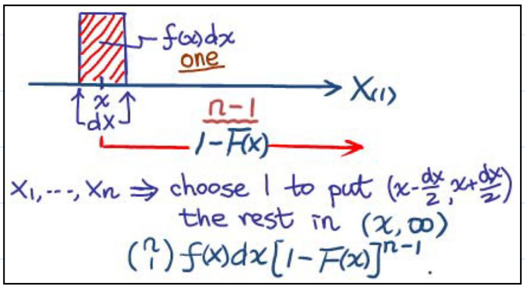
$$f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x)$$
$$= n[1 - F(x)]^{n-1} \left(\frac{d}{dx} F(x)\right) = nf(x)[1 - F(x)]^{n-1}.$$

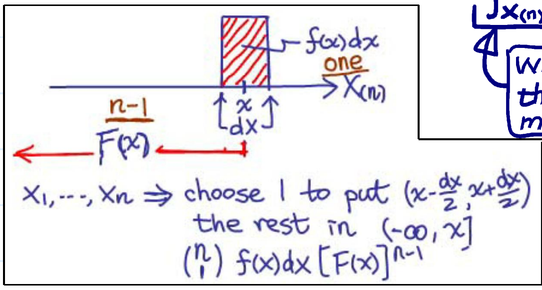
$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x)$$
$$= n[F(x)]^{n-1} \left(\frac{d}{dx} F(x)\right) = nf(x)[F(x)]^{n-1}.$$

- Graphical interpretation for the pdfs of $X_{(1)}$ and $X_{(n)}$.

$n f(x)[1-F(x)]^{n-1} dx$ $\|$ $f_{X_{(1)}}(x)dx$

What does the prob. mean?





$f_{X_{(n)}}(x)dx$

What does the prob. mean?

- Example. n light bulbs are placed in service at time t=0, and allowed to burn continuously. Denote their lifetimes by $X_1, \ldots, X_n$, and suppose that they are i.i.d. with cdf F.

  If burned out bulbs are not replaced, then the room goes dark at time

$$Y = X_{(n)} = \max(X_1, \ldots, X_n).$$

marginal cdf of X_i's

  - If n=5 and F is exponential with λ = 1 per month, then

$$F(x) = 1-e^{-x}, \text{ for } x \geq 0, \text{ and } 0, \text{ for } x < 0.$$

  - The cdf of Y is

$$F_Y(y) = (1-e^{-y})^5, \text{ for } y \geq 0, \text{ and } 0, \text{ for } y < 0,$$

  and its pdf is $5(1-e^{-y})^4 e^{-y}$, for $y \geq 0$, and 0, for $y < 0$.

  - The probability that the room is still lighted after two months is $P(Y > 2) = 1-F_Y(2) = 1-(1-e^{-2})^5$.

➢ Theorem. Suppose that $X_1, \ldots, X_n$ are i.i.d. with pmf p/pdf f. Then, the joint pmf/pdf of $X_{(1)}, \ldots, X_{(n)}$ is

(△)

$$p_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) = n! \times p(x_1) \times \cdots \times p(x_n),$$

or

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) = n! \times f(x_1) \times \cdots \times f(x_n),$$

for $x_1 \leq x_2 \leq \cdots \leq x_n$, and 0 otherwise.

not a cross product set

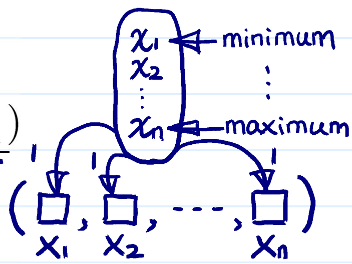
Note: We can derive any marginal distributions of $X_{(1)}, \cdots, X_{(n)}$, or the dist. of the transformations of $X_{(1)}, \cdots, X_{(n)}$ from the joint pmf/pdf.

Why? Check in LNp7-43

Proof. For $x_1 \leq x_2 \leq \cdots \leq x_n$,

$$\begin{aligned} & p_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) \\ &= P(X_{(1)} = x_1, \ldots, X_{(n)} = x_n) \\ &= \sum_{\substack{(i_1,\ldots,i_n): \\ \text{permutations of} \\ (1,\ldots,n)}} P(X_1 = x_{i_1}, \ldots, X_n = x_{i_n}) \\ &\overset{\because \text{ i.i.d.}}{=} \sum_{\substack{(i_1,\ldots,i_n): \\ \text{permutations of} \\ (1,\ldots,n)}} p(x_1) \times \cdots \times p(x_n) \\ &= n! \times p(x_1) \times \cdots \times p(x_n). \end{aligned}$$

symmetric

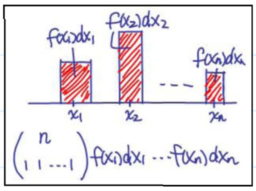


What does the prob. mean?

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1,\ldots,x_n)\,dx_1\cdots dx_n$$
$$\approx P\left(x_1-\tfrac{dx_1}{2}<X_{(1)}<x_1+\tfrac{dx_1}{2},\ldots,\; x_n-\tfrac{dx_n}{2}<X_{(n)}<x_n+\tfrac{dx_n}{2}\right)$$
$$=\sum_{\substack{(i_1,\ldots,i_n):\\ \text{permutations of}\\ (1,\ldots,n)}} P\left(x_{i_1}-\tfrac{dx_{i_1}}{2}<X_1<x_{i_1}+\tfrac{dx_{i_1}}{2},\ldots,\; x_{i_n}-\tfrac{dx_{i_n}}{2}<X_n<x_{i_n}+\tfrac{dx_{i_n}}{2}\right)$$

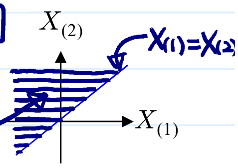
$\because$ i.i.d

$$\approx\sum_{\substack{(i_1,\ldots,i_n):\\ \text{permutations of}\\ (1,\ldots,n)}} f(x_1)\times\cdots\times f(x_n)\,dx_1\cdots dx_n$$
$$= n!\times f(x_1)\times\cdots\times f(x_n)\,dx_1\cdots dx_n.$$



- **Q**: Examine whether $X_{(1)},\ldots,X_{(n)}$ are independent using the Theorem in LNp.7-25.

not a cross product set



➢ Theorem. If $X_1,\ldots,X_n$ are i.i.d. with cdf F and pdf f, then

$X_1,\cdots,X_n$: continuous r.v.'s

1. The pdf of the k^{th} order statistic $X_{(k)}$ is

Note. can be derived from the (Δ) in LNp.7-47 (exercise)

$$f_{X_{(k)}}(x) = \binom{n}{1,k-1,n-k} f(x)F(x)^{k-1}[1-F(x)]^{n-k}.$$

check similar calculation in LNp. 7-42

use $f(x)=\frac{d}{dx}F(x)$ to prove. (exercise)

2. The cdf of $X_{(k)}$ is

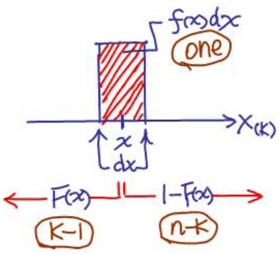
use $F(x)=\int_{-\infty}^{x} f(t)dt$ to prove. (exercise)

$$F_{X_{(k)}}(x) = \sum_{m=k}^{n}\binom{n}{m}[F(x)]^m[1-F(x)]^{n-m}.$$

Proof.

$f_{X_{(k)}}(x)dx$

What does the prob. mean?



$X_1,\cdots,X_n \Rightarrow$ choose 1 to place in $(x-\frac{dx}{2}, x+\frac{dx}{2})$
" $k-1$ " " " $(-\infty, x)$
" $n-k$ " " " (x, ∞)

$$\binom{n}{1\ \ k-1\ \ n-k} f(x)dx\,[F(x)]^{k-1}[1-F(x)]^{n-k}$$

No $\because$ tie

Q: What if $X_1,\cdots,X_n$ are i.i.d. discrete r.v.'s?

YES

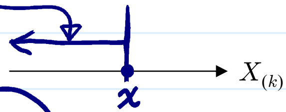
For continuous r.v., $P(X_{(k)}=X_{(k+1)})=0$
For discrete r.v., $P(X_{(k)}=X_{(k+1)})>0$

at least k of $X_1,\cdots,X_n$ fall in $(-\infty,x]$

$$F_{X_{(k)}}(x) = P(X_{(k)}\le x)$$
$$= P(\text{at least } k \text{ of the } X_i\text{'s are} \le x)$$
$$= \sum_{m=k}^{n} P(\text{exact } m \text{ of the } X_i\text{'s are} \le x)$$
$$= \sum_{m=k}^{n}\binom{n}{m}[F(x)]^m[1-F(x)]^{n-m}$$

mutually exclusive

other $n-m$ X_i's are $>x$



➢ Theorem. If $X_1,\ldots,X_n$ are i.i.d. with cdf F and pdf f, then

can obtain using $P(X_{(k)}=x) = F_{X_{(k)}}(x)-F_{X_{(k)}}(x-)$ (exercise)

1. The joint pdf of $X_{(1)}$ and $X_{(n)}$ is

can be derived from (Δ) in LNp.7-47 (exercise)

$$f_{X_{(1)},X_{(n)}}(s,t) = n(n-1)f(s)f(t)[F(t)-F(s)]^{n-2},$$

for $s\le t$, and 0 otherwise.

by the exercise given in LNp.7-32

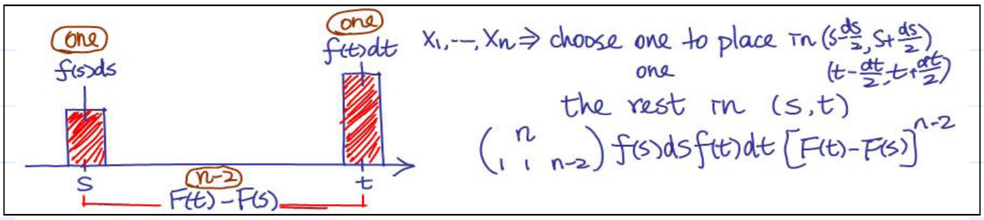
2. The pdf of the range $R = X_{(n)}-X_{(1)}$ is

$$f_R(r) = \int_{-\infty}^{\infty} f_{X_{(1)},X_{(n)}}(u,u+r)\,du,$$

for $r\ge0$, and 0 otherwise.

$f_{X_{(1)},X_{(n)}}(s,t)dsdt$

What does the prob. mean?



➢ Theorem. If $X_1, \ldots, X_n$ are i.i.d. with cdf F and pdf f, then

1. The joint pdf of $X_{(i)}$ and $X_{(j)}$, where $1 \leq i < j \leq n$, is

can be derived from the (△) in LNp.7-47 (exercise)

$$f_{X_{(i)},X_{(j)}}(s,t) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} f(s) f(t) \times [F(s)]^{i-1} [F(t)-F(s)]^{j-i-1} [1-F(t)]^{n-j},$$

for $s \leq t$, and 0 otherwise.

by the exercise given in LNp.7-32

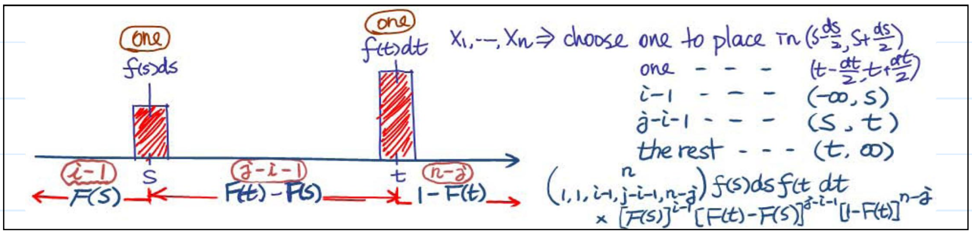
2. The pdf of the j^{th} spacing $S_j = X_{(j)} - X_{(j-1)}$ is

$$f_{S_j}(s) = \int_{-\infty}^{\infty} f_{X_{(j-1)},X_{(j)}}(u, u+s)\, du,$$

for $s \geq 0$, and zero otherwise.

$f_{X_{(i)},X_{(j)}}(s,t)dsdt$

What does the prob. mean?



❖ **Reading**: textbook, Sec 6.3, 6.6, 6.7

Recall Conditional Prob. (LNp.4-1~19)

Conditional Distribution

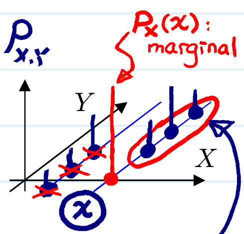
X, Y: maps defined on same Ω

- Definition. Let $\mathbf{X}$ ($\in \mathbb{R}^n$) and $\mathbf{Y}$ ($\in \mathbb{R}^m$) be discrete random vectors and $(\mathbf{X}, \mathbf{Y})$ have a joint pmf $p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})$, then the *conditional joint pmf* of $\mathbf{Y}$ given $\mathbf{X}=\mathbf{x}$ is defined as

notation — still random — fixed

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \equiv P(\underbrace{\{\mathbf{Y}=\mathbf{y}\}}_{B\subset\Omega} | \underbrace{\{\mathbf{X}=\mathbf{x}\}}_{A\subset\Omega}) = \frac{P(\{\mathbf{X}=\mathbf{x}, \mathbf{Y}=\mathbf{y}\})}{P(\{\mathbf{X}=\mathbf{x}\})}$$

$$\boxed{P(B|A) = \frac{P(A\cap B)}{P(A)}} \qquad = \frac{p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})}{p_{\mathbf{X}}(\mathbf{x})} = \frac{\text{joint}}{\text{marginal}}$$



sum of the probabilities, i.e., $P_X(x)$, $= 1$?

if $p_{\mathbf{X}}(\mathbf{x}) > 0$. The probability is defined to be zero if $p_{\mathbf{X}}(\mathbf{x}) = 0$.

➢ Some Notes.

Note. X & Y plays different roles in $P_{Y|X}$

X: not random
Y: random

- For each fixed $\mathbf{x}$, $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$ is a joint pmf for $\mathbf{y}$, since

$$\sum_{\mathbf{y}} p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \frac{1}{p_{\mathbf{X}}(\mathbf{x})} \sum_{\mathbf{y}} p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = \frac{1}{p_{\mathbf{X}}(\mathbf{x})} \times p_{\mathbf{X}}(\mathbf{x}) = 1.$$

- For an event B of $\mathbf{Y}$, the probability that $\mathbf{Y} \in B$ given $\mathbf{X}=\mathbf{x}$ is

$P(Y\in B \,\&\, X=x)/P(X=x)$

$$P(\mathbf{Y} \in B | \mathbf{X} = \mathbf{x}) = \sum_{\mathbf{u}\in B} p_{\mathbf{Y}|\mathbf{X}}(\mathbf{u}|\mathbf{x}). = \left[\sum_{u\in B} P_{X,Y}(x,u)\right] \Big/ P(X=x)$$

- The *conditional joint cdf* of $\mathbf{Y}$ given $\mathbf{X}=\mathbf{x}$ can be similarly defined from the conditional joint pmf $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$, i.e.,

$B=\{Y_1\le y_1, \cdots, Y_m\le y_m\}$

$$F_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = P(\mathbf{Y}\le\mathbf{y}|\mathbf{X}=\mathbf{x}) = \sum_{\mathbf{u}\le\mathbf{y}} p_{\mathbf{Y}|\mathbf{X}}(\mathbf{u}|\mathbf{x}).$$

➢ Theorem. Let $X_1, \ldots, X_m$ be independent and

Poisson$(\lambda_1+\cdots+\lambda_m)$. LNp.7-30

$$X_i \sim \text{Poisson}(\lambda_i),\ i=1, \ldots, m.$$

$\sum_{i=1}^m p_i = \sum_{i=1}^m \frac{\lambda_i}{\lambda_1+\cdots+\lambda_m} = 1$

Let $Y=X_1+\cdots+X_m$, then

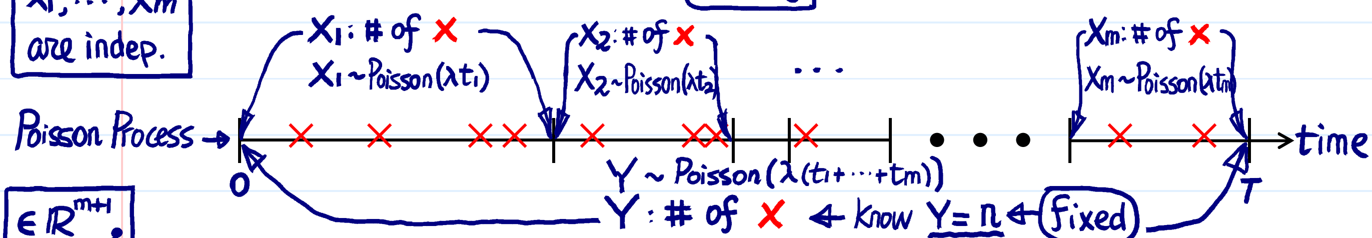
$$(X_1, \ldots, X_m|Y=n) \sim \text{Multinomial}(n, m, p_1, \ldots, p_m),$$

meaning

where $p_i = \lambda_i/(\lambda_1+\cdots+\lambda_m)$ for $i=1, \ldots, m$.

meaning

$X_1, \cdots, X_m$ are indep.



$\in \mathbb{R}^{m+1}$, but the dim. of its pts is only m

Proof. The joint pmf of $(X_1, \ldots, X_m, Y)$ is

$$p_{\mathbf{X},Y}(x_1,\ldots,x_m,n) = P(\{X_1=x_1,\ldots,X_m=x_m\}\cap\{Y=n\})$$

$$= \begin{cases} P(X_1=x_1,\ldots,X_m=x_m), & \text{if } x_1+\cdots+x_m=n, \\ 0, & \text{if } x_1+\cdots+x_m\ne n. \end{cases}$$

Furthermore, the distribution of Y is Poisson$(\lambda_1+\cdots+\lambda_m)$, i.e., LNp.7-30

$$p_Y(n) = P(Y=n) = \frac{e^{-(\lambda_1+\cdots+\lambda_m)}(\lambda_1+\cdots+\lambda_m)^n}{n!}.$$

Therefore, for $\mathbf{x}=(x_1, \ldots, x_m)$ wheres $x_i\in\{0,1,2, \ldots\}$, $i=1, \ldots, m$, and $x_1+\cdots+x_m=n$, the conditional joint pmf of $\mathbf{X}$ given $Y=n$ is

range of $\mathbf{X}|Y=n$

$0\le x_1\le n$ ⋮ $0\le x_m\le n$

$$p_{\mathbf{X}|Y}(\mathbf{x}|n) = \frac{p_{\mathbf{X},Y}(x_1,\ldots,x_m,n)}{p_Y(n)} = \frac{\prod_{i=1}^m \frac{\cancel{e^{-\lambda_i}}\lambda_i^{x_i}}{x_i!}}{\frac{\cancel{e^{-(\lambda_1+\cdots+\lambda_m)}}(\lambda_1+\cdots+\lambda_m)^{n = x_1+\cdots+x_n}}{n!}}$$

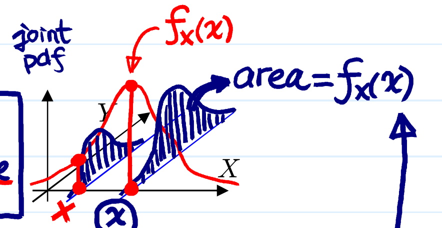
$$= \frac{n!}{x_1!\times\cdots\times x_m!}\times\left(\underbrace{\frac{\lambda_1}{\lambda_1+\cdots+\lambda_m}}_{p_1}\right)^{x_1}\times\cdots\times\left(\underbrace{\frac{\lambda_m}{\lambda_1+\cdots+\lambda_m}}_{p_m}\right)^{x_m}.$$

$p_1+\cdots+p_m=1$.

- Definition. Let $\mathbf{X}$ $(\in\mathbb{R}^n)$ and $\mathbf{Y}$ $(\in\mathbb{R}^m)$ be continuous random vectors and $(\mathbf{X}, \mathbf{Y})$ have a joint pdf $f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})$, then the *conditional joint pdf* of $\mathbf{Y}$ given $\mathbf{X}=\mathbf{x}$ is defined as

random fixed

maps defined on same Ω



notation

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \equiv \frac{f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X}}(\mathbf{x})} = \frac{\text{joint}}{\text{marginal}},$$

if $f_{\mathbf{X}}(\mathbf{x})>0$, and 0 otherwise.

$\text{area} = \int_{-\infty}^{\infty} f(x,y)\,dy = f_X(x) \overset{?}{\ne} 1$

➤ Some Notes.

- P(**X**=**x**)=0 for a continuous random vector **X**.
- The justification of $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$ comes from

$\{x_1-\frac{\Delta x_1}{2}<X_1\le x_1+\frac{\Delta x_1}{2},\ x_2-\frac{\Delta x_2}{2}<X_2\le x_2+\frac{\Delta x_2}{2},\ \ldots,\ x_n-\frac{\Delta x_n}{2}<X_n\le x_n+\frac{\Delta x_n}{2}\}$

joint cdf of $Y|X\in x\pm\Delta X/2$

$\{Y_1\le y_1, Y_2\le y_2, \ldots, Y_m\le y_m\}$

$$P(\mathbf{Y}\le\mathbf{y}|\underbrace{\mathbf{x}-(\Delta\mathbf{x}/2)<\mathbf{X}\le\mathbf{x}+(\Delta\mathbf{x}/2)}_{A})$$

$$=\frac{\boxed{\int_{-\infty}^{\mathbf{y}}}\boxed{\int_{\mathbf{x}-(\Delta\mathbf{x}/2)}^{\mathbf{x}+(\Delta\mathbf{x}/2)}}f_{\mathbf{X},\mathbf{Y}}(\mathbf{u},\mathbf{v})\,d\mathbf{u}d\mathbf{v}}{\int_{\mathbf{x}-(\Delta\mathbf{x}/2)}^{\mathbf{x}+(\Delta\mathbf{x}/2)}f_{\mathbf{X}}(\mathbf{t})\,d\mathbf{t}}\begin{matrix}\approx P(A\cap B)\\ \approx P(A)\end{matrix}$$

becomes "=" when $\Delta X\to 0$

$$\approx\frac{\int_{-\infty}^{\mathbf{y}}f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{v})|\cancel{\Delta\mathbf{x}}|\,d\mathbf{v}}{f_{\mathbf{X}}(\mathbf{x})|\cancel{\Delta\mathbf{x}}|}=\int_{-\infty}^{\mathbf{y}}\frac{f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{v})}{f_{\mathbf{X}}(\mathbf{x})}\,d\mathbf{v}$$

joint pdf of $Y|X=x$ by $\frac{\partial^m}{\partial y_1\cdots\partial y_m}F_{Y|X}$

- For each fixed **x**, $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$ is a joint pdf for **y**, since

Note x & y play different roles

$$\int_{-\infty}^{\infty}f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})\,d\mathbf{y}=\frac{1}{f_{\mathbf{X}}(\mathbf{x})}\int_{-\infty}^{\infty}f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})\,d\mathbf{y}=\frac{1}{f_{\mathbf{X}}(\mathbf{x})}\times f_{\mathbf{X}}(\mathbf{x})=1.$$

$\int_{\mathbb{R}^m}$

- For an event B of **Y**, we can write

$$P(\mathbf{Y}\in B|\mathbf{X}=\mathbf{x})=\int_B f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})\,d\mathbf{y}.$$

- The *conditional joint cdf* of **Y** given **X**=**x** can be similarly defined from the conditional joint pdf $f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$, i.e.,

$$F_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})=P(\mathbf{Y}\le\mathbf{y}|\mathbf{X}=\mathbf{x})=\int_{-\infty}^{\mathbf{y}}f_{\mathbf{Y}|\mathbf{X}}(\mathbf{t}|\mathbf{x})\,d\mathbf{t}.$$