$(\Omega, \mathcal{A}, P)$ - a probability space.

$\underset{\sim}{X} = (X_1, \ldots, X_n)$ is a random vector.

$\Leftrightarrow X_1, \ldots, X_n$ are real random variables.

$$\underset{\sim}{X} : \Omega \longrightarrow \mathbb{R}^d$$

Joint probability distribution $\underset{\sim}{X}$:

$$P_{\underset{\sim}{X}}(B) = P(X \in B), \quad B \subset \mathbb{R}^k$$

$$F_{\underset{\sim}{X}}(\underset{\sim}{a}) = P(X_1 \leq a_1, \ldots, X_k \leq a_k),$$

$$\underset{\sim}{a} = (a_1, \ldots, a_k)$$

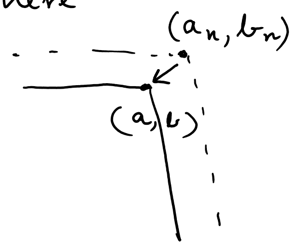
$$\therefore F_{\underset{\sim}{X}} : \mathbb{R}^k \longrightarrow \mathbb{R}.$$

Eg: $k=2$.

* F is "right-continuous" everywhere

[ie, $a_n \searrow a$, $b_n \searrow b$, then

$$F(a_n, b_n) \longrightarrow F(a, b).$$



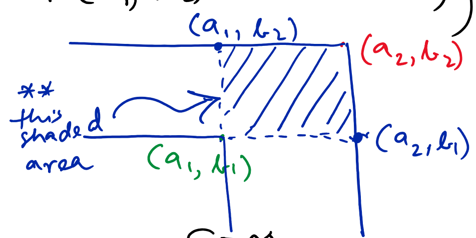
** $\Delta F_{(a_1, b_1), (a_2, b_2)} = \Big(F(a_2, b_2) - F(a_2, b_1) - F(a_1, b_2) + F(a_1, b_1) \Big)$

$a_1 \leq a_2, \quad b_1 \leq b_2$

(Stronger than "non-decreasing" property)

≥ 0

$\forall (a_1, b_1), (a_2, b_2)$.



* $F(a, b) \longrightarrow \begin{cases} 0 \\ 1 \end{cases}$ as $a \wedge b \longrightarrow \begin{cases} -\infty \\ \infty \end{cases}$

$a \wedge b$ ↳ $\min\{a, b\}$

$a \wedge b \longrightarrow -\infty$

$\Rightarrow$ atleast one of $a, b \rightarrow -\infty$

& $a \wedge b \longrightarrow \infty$

$\Rightarrow$ atleast one of $a, b \to \infty$
& , $a \wedge b \longrightarrow \infty$
$\Rightarrow$ both $a, b \longrightarrow \infty$.

Important: In 2 or higher dimensions,
F is continuous everywhere
$\Updownarrow$
$P(X=a) = P(Y=b) = 0 \ \forall\ a, b \in \mathbb{R}$.
or, F has uncountably many discontinuities.

Marginals:

$$F_X(a) = \lim_{b \to \infty} F(a, b).$$

$$F_Y(b) = \lim_{a \to \infty} F(a, b).$$

Special cases:

① (X,Y) - discrete if $\exists$ countable $D \subset \mathbb{R}^2$
s.t. $P((X,Y) \in D) = 1$

$p(x,y) = P(X=x, Y=y)$, $(x,y) \in \mathbb{R}^2$; joint pmf
$p(x,y) = 0$ if $(x,y) \notin D$ [$\to$ ensures that $p(x,y) > 0 \ \forall\ (x,y) \in D$.]

$$P_{X,Y}(B) = \sum_{(x,y) \in D \cap B} p(x,y)$$

$$p_X(x) = \sum_y p(x,y), \quad p_Y(y) = \sum_x p(x,y) \quad \leftarrow \text{Marginals}$$

② (X,Y) - jointly (absolutely) continuous if $\exists\ f \geq 0$ on $\mathbb{R}^2$
s.t. $P_{X,Y}(B) = \iint_B f(x,y)\,dx\,dy$, $B \subset \mathbb{R}^2$ (borel)

$$F_{X,Y}(a,b) := \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y)\,dx\,dy = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)\,dy\,dx$$

[Fubini's theorem
— gives the conditions

[Fubini's theorem – gives the conditions for interchanging the "$\int$"s or "$\sum$"s.]

$$F_X(a) = \int_{-\infty}^{\infty}\int_{-\infty}^{a} f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{a}\Big(\underbrace{\int_{-\infty}^{\infty} f(x,y)\,dy}_{=\,f_X(x)}\Big)dx = \int_{-\infty}^{a}\Big(\quad\Big)\cdot dx$$

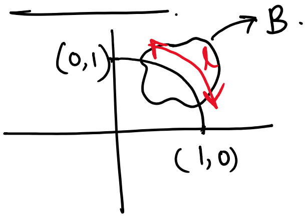
$\to f^n$ of X.

$f_X(x) \leftarrow$ density of X. (why? ↑)

ie, "integrating out" y gives density of X.

* X, Y absolutely continuous $\not\Rightarrow$ (X,Y) jointly (abs) continuous.

Picture: 1



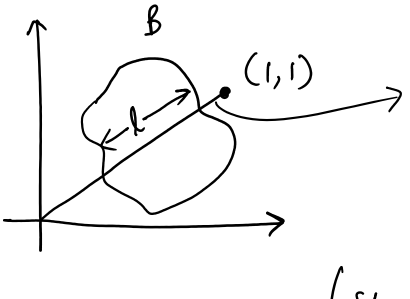
Arc from unit circle.

$0 \le x \le 1$,
$0 \le y \le 1$.

ie, entire probability mass lies on the arc.

$$P_{X,Y}(B) = \frac{\text{Arc length}(l)}{\pi/2}$$

Picture 2:



entire probability mass distributed on this line.

Here, $P_{X,Y}(B) = \left(\frac{l}{\sqrt{2}}\right)$

(show that, both the marginals are Unif$(0,1)$, but they do not have a joint density.)

Special result:

(X,Y) has a density f on I → open region in $\mathbb{R}^2$.

(X, Y) has a density f ... $\hookrightarrow$ open region in $\mathbb{R}^2$.

$h: I \longrightarrow \tilde{I}$ (an open region in $\mathbb{R}^2$.)
(bijection)

$g = h^{-1}: \tilde{I} \longrightarrow I$.

$h: (x, y) \longmapsto (u, v)$

$g: (u, v) \longmapsto (x, y)$

$\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}, \frac{\partial y}{\partial u}, \frac{\partial y}{\partial v}$ exist and are continuous in I.

The Jacobian, $J(u,v) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$

Suppose, $\det(J(u,v)) \neq 0$ on $\tilde{I}$.

Under all these conditions:

$$f_{U,V}(u,v) = f_{X,Y}(g(u,v)) \cdot |\det(J(u,v))| \quad \forall (u,v) \in \tilde{I}$$

. . . .

$\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)$ has density $f_{\underset{\sim}{X}}$ on $I \leftarrow$ open region in $\mathbb{R}^k$.

$h: I \longrightarrow \tilde{I} \subset \mathbb{R}^k$

$g = h^{-1}: \tilde{I} \longrightarrow I$.

$(y_1, \ldots, y_k) \longmapsto (x_1, \ldots, x_k)$.

$J(\underset{\sim}{y}) = \left(\left(\frac{\partial x_i}{\partial y_j} \right) \right)_{k \times k}$ exists and continuous on I,

& $\det(J(\underset{\sim}{y})) \neq 0$ on I.

then $\underset{\sim}{Y} = h(\underset{\sim}{X})$ has density $f_{\underset{\sim}{Y}}(y_1, \ldots, y_k) = f_{\underset{\sim}{X}}(g(y))$.

Special Case:

Let A be a non-singular matrix

$\underset{\sim}{\beta}$ be a k-dimensional column vector.

$h(\underset{\sim}{x}) = A\underset{\sim}{x} + \underset{\sim}{\beta} \longleftarrow$ "translation"

$g(\underset{\sim}{y}) = A^{-1}(\underset{\sim}{y} - \underset{\sim}{\beta})$

$$g(\underset{\sim}{y}) = A^{-1}(\underset{\sim}{y} - \underset{\sim}{\beta})$$

$$\det(J(\underset{\sim}{y})) = \det(A^{-1}).$$

$\Rightarrow \underset{\sim}{Y} = A\underset{\sim}{X} + \underset{\sim}{\beta}$ has density

$$f_Y(\underset{\sim}{y}) = \frac{1}{|\det A|} \cdot f_{\underset{\sim}{X}}(A^{-1}(\underset{\sim}{y} - \underset{\sim}{\beta}))$$

A <u>general approach</u> if this method <u>fails</u>:

$\underset{\sim}{X}$ has density $f_{\underset{\sim}{X}}$ on I.

$\underset{\sim}{Y} = h(\underset{\sim}{X})$. "Special case" doesn't hold.

$h: I \longrightarrow \tilde{I}$.

Fix a point $\underset{\sim}{y} \in \tilde{I}$, take $\underset{\sim}{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n) > 0$ (ie, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, $\vdots$, $\varepsilon_n > 0$)

small enough so that —

$$[\underset{\sim}{y}, \underset{\sim}{y} + \varepsilon) \subset \tilde{I}$$

$$\shortparallel$$

$$[y_1, y_1 + \varepsilon) \times \dots \times [y_k, y_k + \varepsilon).$$

<u>Compute</u>:

$$P(\underset{\sim}{Y} \in [\underset{\sim}{y}, \underset{\sim}{y} + \varepsilon)) = P(h(\underset{\sim}{X}) \in [\underset{\sim}{y}, \underset{\sim}{y} + \varepsilon))$$

look at $\underbrace{\frac{1}{\left(\prod_{i=1}^{n} \varepsilon_i\right)}}_{\text{volume of that box.}} \cdot P(\underset{\sim}{Y} \in [\underset{\sim}{y}, \underset{\sim}{y} + \varepsilon))$

"Lebesgue's Differentiation Theorem"

now, let $\varepsilon_i \longrightarrow 0$ & see if there exists a limit.

Conditional Distribution:

X, Y — random variables.

(X, Y) has a joint distribution.

How to define: Conditional Distribution of Y given $X = x$.

<u>Aim</u>: to define $Q(x, B) \overset{??}{=\!=} P(Y \in B \mid X = x)$

Q → a function; B → set; x → point.

ie, both a set & point f^n.

Such that, Q must satisfy:

① for each $x \in \mathbb{R}$, $Q(x,B)$ is a prob. on $\mathbb{R}$.

② for each $B \subset \mathbb{R}$, $Q(x,B)$ is measurable f^n of x.

③ For every pair of borel sets $A, B \subset \mathbb{R}$.

$$P(X \in A, Y \in B) = E(\underbrace{Q(X,B)}_{\text{A random variable: } P(Y\in B \mid X)} \cdot 1_A(X)).$$

does the job of keeping $X \in A$.

A random variable: $P(Y \in B \mid X)$.

Cases where we can identify a legitimate "Q"?

Case-1: X- discrete.

Case-2: When (X,Y) has a joint density

then, $h(x,y) = \begin{cases} \dfrac{f_{X,Y}(x,y)}{f_X(x)} & \forall x, \text{ s.t } f_X(x) > 0 \\ \dfrac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}y^2} & \text{otherwise} \end{cases}$

can be anything else as well !!

for these 2 cases above —

$$\therefore Q(x,B) = \int_B h(x,y)\,dy.$$

[Exercise: prove that this specific $Q(x,B)$ follows the three conditions.]

Case-3: X, Z - independent r.vs. → necessary.

$$Y := h(X,Z)$$

$$Q(x,B) = P(h(x,Z) \in B)$$

[if x is fixed, $h(x,Z)$ is a random variable]

Claim:

Q is a candidate for $P(Y \in B \mid X = x)$.

Exercise: prove the claim. (hint: starting from very def^n of independence)

Why does this makes sense intuitively?

$$P(h(X,Z) \in B \mid X = x) = P(h(x,Z) \in B)$$

[$\because X, Z$ are independent, the condition that $X = x$

[$\because X, Z$ are independent,
the condition that $X = x$
only affects X, not Z.]