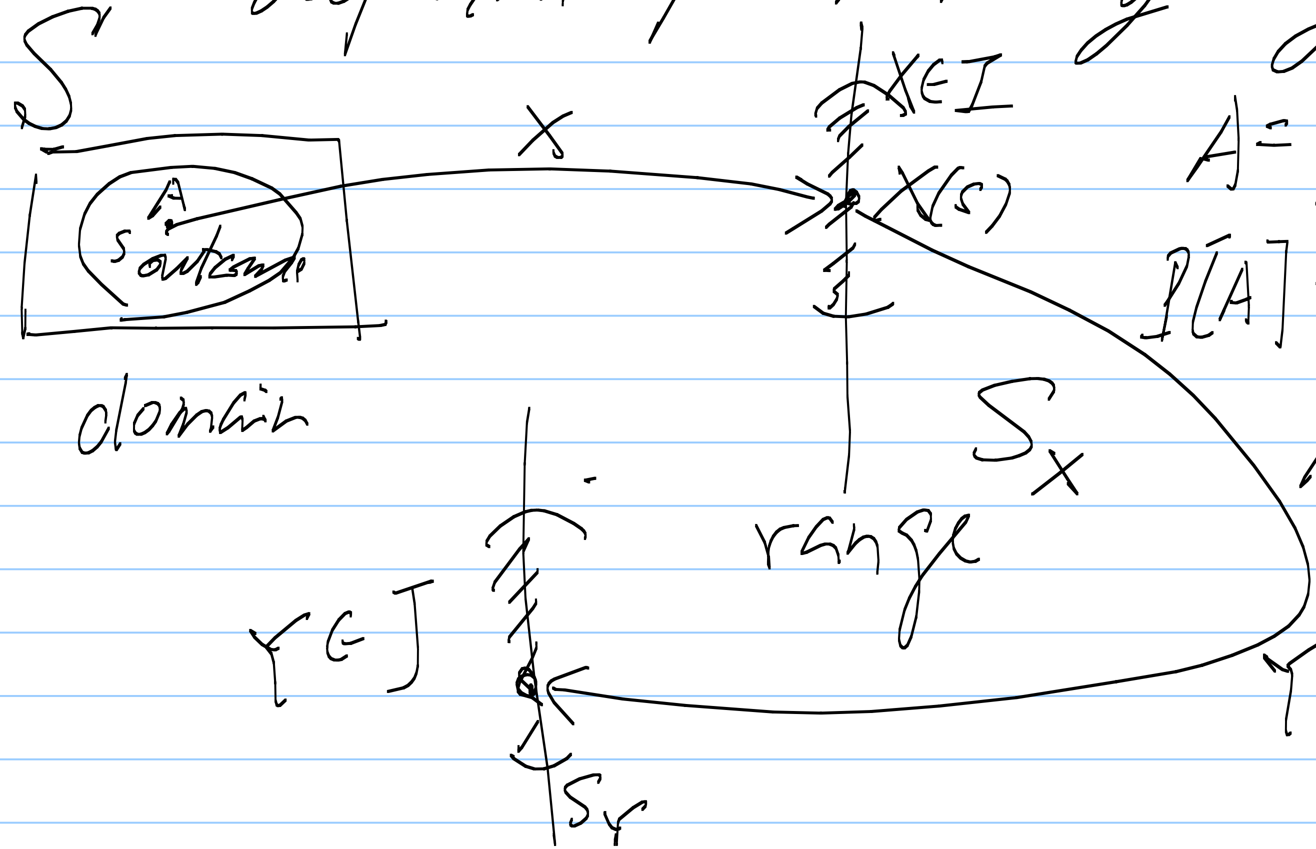
CHAP 4

Recall: One Random Variable is defined pictorially by



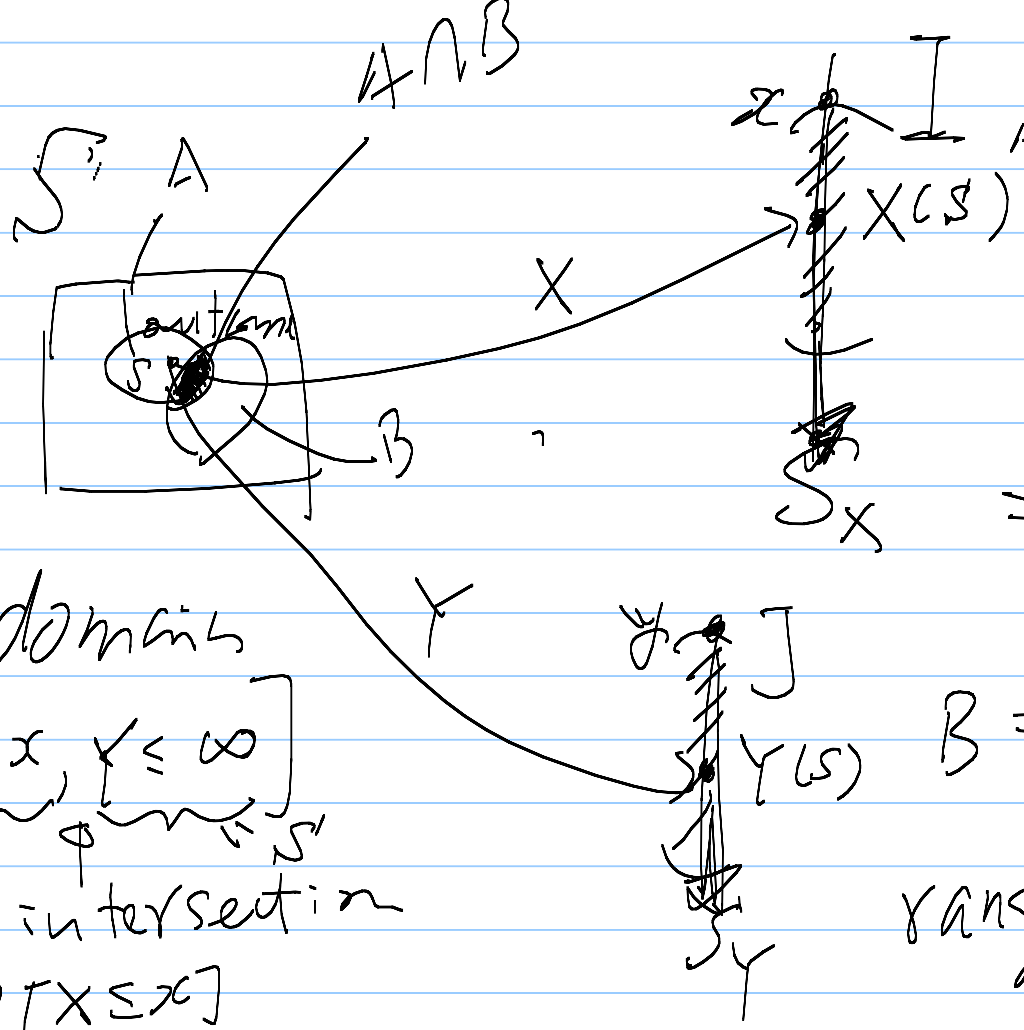
$A = \{ X \in I \}$

$P[A] = P[X \in I]$

$A = \{ Y \in J \}$

$Y = g(X)$

Define Two Random Variables Pictorially:



$A = \{X \in I\}$

$P[A \cap B]$

$= P[X \in I, Y \in J]$ intersection

$B = \{Y \in J\}$

$$\boxed{P[\underbrace{X \le x}_{A}, \underbrace{Y \le -\infty}_{\Phi}] = A \cap \Phi = P[\phi] = 0}$$

domain

$$P[\underbrace{X \le x}_{\phi}, \underbrace{Y \le \infty}_{= S'}]$$

$=$ intersection

$P[X \le x]$

range $= \{(x, y) \mid x \in S_X, y \in S_Y\}$

$= S_{X,Y}$

$$F_{X,Y}(x,y) \triangleq P[X \le x, Y \le y], \quad \forall x, y$$

Note: $F_{X,Y}(x,y)$ is continuous from the right in x and y.

If we consider $F_{X,Y}(x,\infty)$,

$$\underbrace{F_{X,Y}(x,\infty)} = P[\underbrace{X \le x}_{\{X \le x\}} \cap \underbrace{Y \le \infty}_{= S}]$$

$$= P[X \le x] = F_X(x)$$

$F_X(x)$ is called the marginal CDF of X.

$F_{X,Y}(x,y)$ is called the joint CDF of X and Y.

$F_{X,Y}(x,y)$ is sufficient to completely describe the probability model of

X and Y,

pf: Given $\{X \le x\} \subset \{X \le x_1\}$ since $x \le x_1$

$\{Y \le y\} \subset \{Y \le y_1\}$ since $y \le y_1$,

$$\Rightarrow F_{X,Y}(x,y) = P[\underbrace{X \le x}_{A}, \underbrace{Y \le y}_{B}] \le P[X \le x_1, Y \le y_1] = F_{X,Y}(x_1, y_1)$$

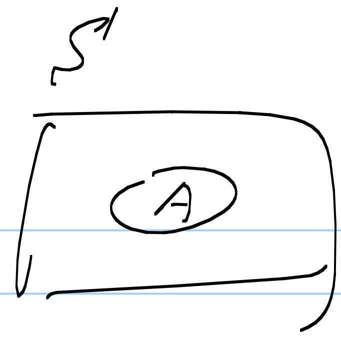
since $A \cap B \subset \{X \le x_1\} \cap \{Y \le y_1\}$,

(f) $F_{X,Y}(\infty, \infty) = 1$,

pf: $F_{X,Y}(\infty,\infty) = P[X \le \infty, Y \le \infty] = P[S \cap S] = P[S]$

$= 1$

$$\mathbb{P}[\underbrace{(X, Y) \in B}_{\text{an event}}] = \mathbb{P}[B]$$

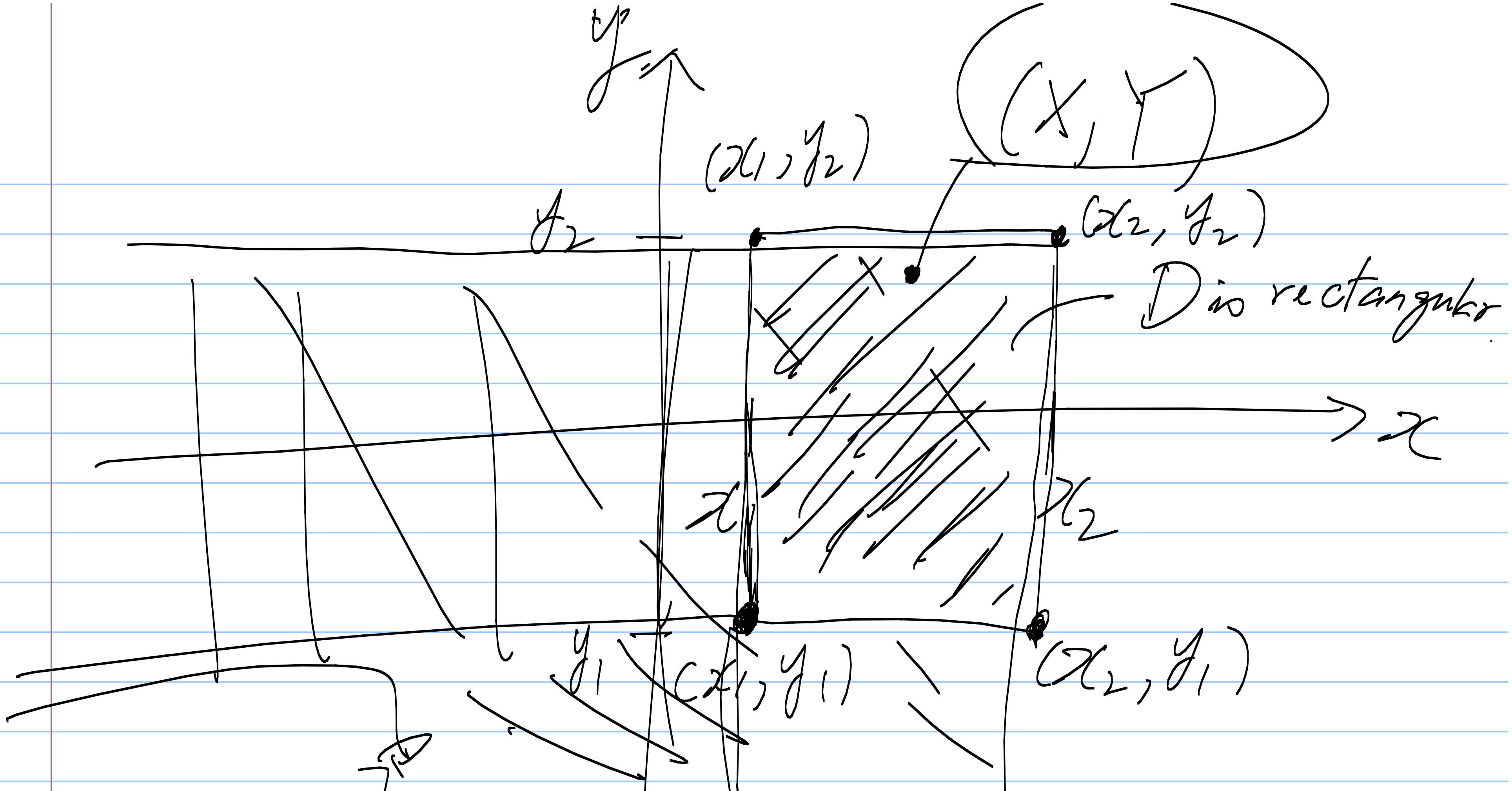


$$\mathbb{P}_X(x) \stackrel{\circ}{=} \mathbb{P}[\underbrace{X = x}_{A}] = \mathbb{P}[A \cap \underbrace{S'}_{Y \in S_Y}]$$

$$= \mathbb{P}[X = x, Y \in S_Y]$$

$$= \sum_{y \in S_Y} \mathbb{P}_{X,Y}(x, y) \quad \text{Q.E.D.}$$

Defn: X and Y are called two continuous random variables if their joint CDF $F_{X,Y}(x, y)$ is continuous in x and y, and if the derivative $\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$ exists.



$$P[(X,Y)\in D] = F_X(x_2,y_2) - F_X(x_1,y_2) - F_X(x_2,y_1) + F_X(x_1,y_1)$$

$$P[A] \equiv P[(X, Y) \in A]$$

2-dimensional region

$$= \iint_A f_{X,Y}(x, y)\, dx\, dy$$

$$\boxed{F_X(x)} \triangleq P[X \le x] = P[X \le x, Y \le \infty]$$

$$= \int_{-\infty}^{x} \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(u, v)\, dv}_{= f_X(u)}\, du$$

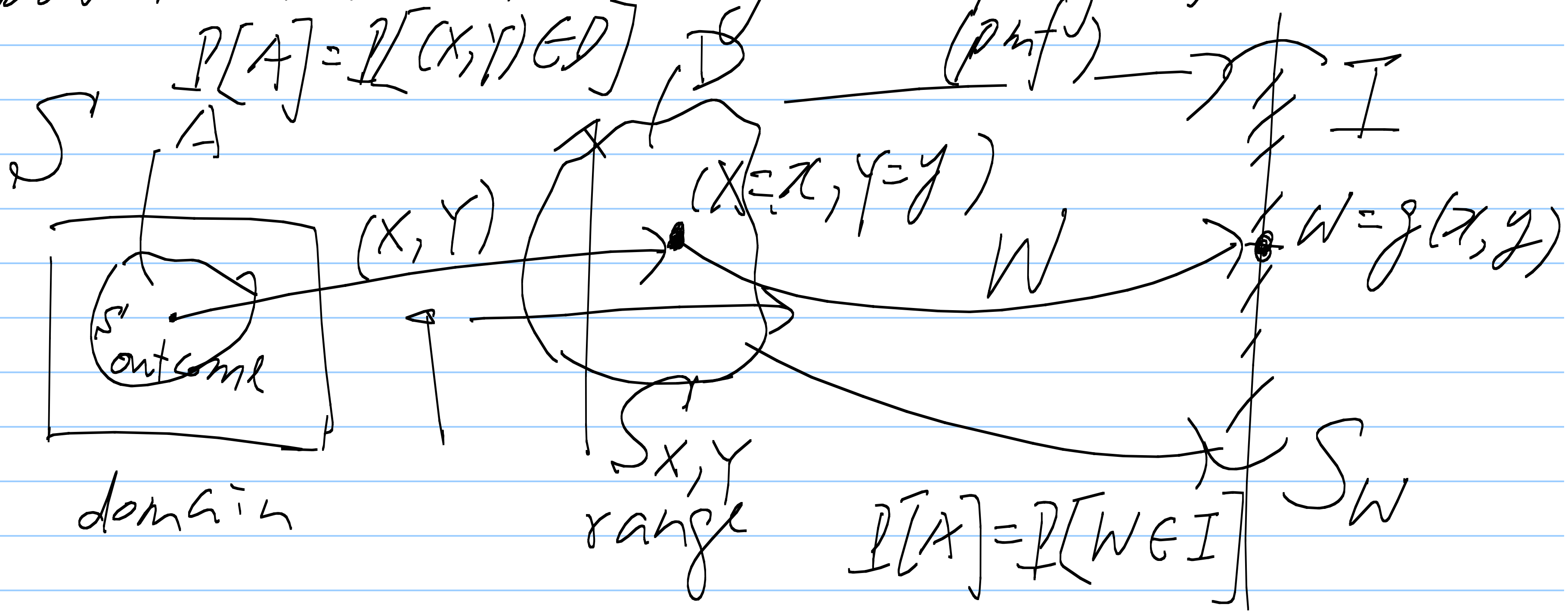
marginal CDF of Y

marginal PDF of X

Q.E.D.

$$W = g(X, Y)$$

We want to statistically characterize the derived random variable $W = g(X, Y)$ provided with the joint pdf of X and Y.



Thus, $P[W \in I] = P[(X,Y) \in D]$.

* If X and Y are discrete and if $I = \{w\}$, then

$$P[W \in I] = P[W = w] = P_W(w) = P[g(X,Y) = w]$$
$$= \sum_{\substack{x \quad y \\ g(x,y)=w}} \sum P_{X,Y}(x,y)$$

* If X and Y are continuous, and if W is also continuous, then ($I = \{x \mid x \leq w\}$)

$$P[W \in I] = P[W \leq w] = F_W(w) = P[g(X,Y) \leq w]$$
$$= \iint_{g(x,y) \leq w} f_{X,Y}(x,y)\, dx\, dy$$

* Define $W = \max(X, Y)$. Then,

$$F_W(w) \triangleq P[W \le w] = P[\max(X, Y) \le w]$$

$$= P[X \le w, Y \le w]$$

and

$$= \int_{-\infty}^{w}\int_{-\infty}^{w} f_{X,Y}(x, y)\, dx\, dy$$

Q.E.D.

$$F_W(w) = \frac{w^2}{15}[u(w) - u(w-3)] + \frac{w}{5}[u(w-3) - u(w-5)] + u(w-5)$$

$$f_W(w) = \frac{2w}{15}[u(w) - u(w-3)] + \frac{1}{5}(u(w-3) - u(w-5)) \Leftarrow$$

By law of unconscious statistician,

$$E[W] = E[g(X,Y)] = (*)$$

(i) If X and Y are discrete,

$$(*) = \sum_x \sum_y g(x,y)\, P_{X,Y}(x,y)$$

(ii) If X and Y are continuous,

$$(*) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\, dx\, dy$$

$$\sigma_{X,Y}^2 \equiv Cov[X, Y]$$

$$= E[(X-\mu_X)(Y-\mu_Y)]$$

summation

$$E[X+Y] = E[X] + E[Y]$$

$$\boxed{E[XY] = r_{X,Y}}$$

If $Cov[X, Y] = 0$, then

$$\underbrace{E[XY]}_{r_{X,Y}} = \mu_X \mu_Y \Leftrightarrow E[XY] = E[X]E[Y]$$

product

Notes: ① $E[X+Y] = E[X] + E[Y]$ for all X & Y,

and

② $E[XY] \neq E[X]E[Y]$ in general.

But, if X and Y are uncorrelated, then

$$E[XY] = E[X]\,E[Y].$$

$$Cov[X, Y] = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$

$$\rho_{X,Y} = E\left[\frac{(X-\mu_X)}{\sigma_X}\frac{(Y-\mu_Y)}{\sigma_Y}\right] \equiv \text{correlation coefficient}$$

We used to call

X and Y are positively correlated if $1 \geq \rho_{X,Y} > 0$

negatively correlated if $-1 \leq \rho_{X,Y} < 0$

uncorrelated if $\rho_{X,Y} = 0$

Notes: ① $X = Y$ means "X and Y are equivalent everywhere", i.e., $f_X(x) = f_Y(x)$, $\forall x$.

② $X \overset{MS}{=} Y$ means "X and Y are equivalent in the mean-square sense", i.e.,

$$E[(X-Y)^2] = 0$$

③ $X \overset{P}{=} Y$ means "X and Y are equivalent with probability one" or, almost everywhere",

i.e., $P[X = Y] = 1$.

$$P_{X,Y|B}(x, y) \overset{\circ}{=} P[X = x, Y = y \mid B] = \frac{P[\{X = x\} \cap \{Y = y\} \cap B]}{P[B]}$$

$P[(X, Y) \in B]$

#

$$\underline{f_{X,Y|B}}(x,y) \triangleq \lim_{\substack{dx\to 0\\ dy\to 0}} P\left[\underbrace{x \le X < x+dx}, \underbrace{y \le Y < y+dy} \,\middle|\, B\right] \Big/ dx\,dy$$

$$= \lim_{\substack{dx\to 0\\ dy\to 0}} \frac{P\left[\{x \le X < x+dx\} \cap \{y \le Y < y+dy\} \cap B\right]}{P[B]\ dx\ dy}$$

As $dx \to 0$ and $dy \to 0$,

$$\underbrace{f_{X,Y|B}(x,y)\,\underline{dx\,dy}} = P\left[x \le X < x+dx,\ y \le Y < y+dy \,\middle|\, B\right]$$

$f_{X,Y|B}(x, y)$

$B = \{W = w\}$ if W is discrete

or

$B = \{w \le W < w + dw\}$ for any

typed random variable

W and for $dx \to 0$.

$$P[\underbrace{x \le X < x+dx}_{A} \mid \underbrace{y \le Y < y+dy}_{B}] \checkmark$$

$$= \frac{P[x \le X < x+dx,\ y \le Y < y+dy]}{P[y \le Y < y+dy]}$$

As $dx \to 0$ and $dy \to 0$,

$$f_{X,Y}(x,y)\,\cancel{dx\,dy} = \underbrace{f_{X|Y}(x|y)}\,\cancel{dx}\ \underbrace{f_Y(y)}\,\cancel{dy}$$

$$\Rightarrow f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \#$$

$\boxed{E[g(X,Y) \mid Y]}$ represents a function of Y, which is a random variable.

$E[g(X,Y) \mid Y=y]$ represents a function of y, which is deterministic.

(If Y is continuous, the conditioning event $\{Y=y\} = \{y \le Y < y + dy\}$)

If Y is continuous,

$$E[g(X)] = E\left[E[g(X) \mid Y]\right]$$

$$= \int_{-\infty}^{\infty} E[g(X) \mid Y=y]\, f_Y(y)\, dy$$

If Y is discrete,

$$E[g(X)] = \sum_{y \in S_Y} E[g(X) \mid Y=y]\, P_Y(y)$$

$$* \quad E[g(X,Y)] = E\left[E[g(X,Y) \mid X]\right]$$

$$\left(E[X_1] = E[E[E[\cdots E[X_1 \mid X_n] \mid X_{n-1}] \cdots \mid X_2]] \right)$$

$$= E\left[E[g(X,Y) \mid Y]\right]$$

for any function g, any rv's X and Y.

* Recall that two events A and B are independent iff (if and only if)

$$P[A \cap B] = P[A]\,P[B] \iff P[A \mid B] = P[A] \quad \checkmark$$

* Defn: Two discrete random variables X and Y are called (mutually) independent iff

$$P[X=x, Y=y] = P[X=x]\,P[Y=y] \text{ for all } x \text{ and } y.$$

Two continuous random variables X and Y are called independent iff

$$P[x \le X < x+dx,\ y \le Y < y+dy]$$
$$= \underbrace{P[x \le X < x+dx]}\ \underbrace{P[y \le Y < y+dy]}$$

for all x, y, as $dx \to 0$ and $dy \to 0$.

As $dx \to 0$ and $dy \to 0$,

$$f_{X,Y}(x,y)\,\cancel{dx\,dy} = f_X(x)\,\cancel{dx}\,f_Y(y)\,\cancel{dy}$$

$$\Rightarrow \boxed{f_{X,Y}(x,y) = f_X(x)\,f_Y(y)}.$$

If X and Y are independent,

then $f_{X|Y}(x|y) = f_X(x)$ if X and Y are continuous

$P_{X|Y}(x|y) = P_X(x)$ if X and Y are discrete.

For any two random variables X and Y, we know that

1. $E[X+Y] = E[X] + E[Y]$
2. $E[XY] \neq E[X]E[Y]$ in general
3. $\boxed{E[XY] = E[X]E[Y]}$ iff X and Y are uncorrelated
4. uncorrelated $\Longleftarrow$ independent

uncorrelated $\not\Longrightarrow$ independent in general

uncorrelated $\Longrightarrow$ independent (bivariate Gaussian)

Independence is sufficient for uncorrelation.

Uncorrelation is necessary for independence.

* $\tilde{\mu}_2(x) = \mu_2 + \rho \dfrac{\sigma_2}{\sigma_1}(x - \mu_1)$ is the conditional expectation of Y given $X = x$.

$\tilde{\sigma}_2^2 = \left(\sigma_2\sqrt{1-\rho^2}\right)^2$ is the conditional variance of Y given $X = x$.

For bivariate Gaussian X and Y,

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\} \leftarrow f_X(x)$$

$$\cdot \left(\frac{1}{\sqrt{2\pi\hat{\sigma}_2^2}} \exp\left\{-\frac{(y-\hat{\mu}_2(x))^2}{2\hat{\sigma}_2^2}\right\}\right)$$

$$f_{Y|X}(y|x)$$

You can also write

$$f_{X,Y}(x,y) = f_Y(y)\, f_{X|Y}(x|y)$$

where $f_Y(y) \sim N[\mu_2, \sigma_2^2]$

$f_X(x) \sim N[\mu_1, \sigma_1^2]$

and $f_{X|Y}(x|y) \sim N[\tilde{\mu}_1(y), \tilde{\sigma}_1^2]$

with $\tilde{\sigma}_1^2 = \sigma_1^2(1-\rho^2)$

$\tilde{\mu}_1(y) = \mu_1 + \rho \dfrac{\sigma_1}{\sigma_2}(y - \mu_2)$

If $\rho = 0$, i.e., X and Y are uncorrelated, then

$$f_{X,Y}(x,y) = \underbrace{\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}}_{f_X(x)} \cdot \underbrace{\frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}}_{f_Y(y)}$$

this indicates that X and Y are independent.

For bivariate Gaussian X and Y,

$$Var[Y \mid X=x] = \sigma_2^2(1-\rho^2) \le \sigma_2^2 = Var[Y]$$

$$Var[X \mid Y=y] = \sigma_1^2(1-\rho^2) \le \sigma_1^2 = Var[X]$$

#

$F_{X,Y}(x, y)$ is nondecreasing in arguments x and y.

Define $W = \min(X, Y)$.

Then,

$$F_W(w) \triangleq P[W \le w] = P[\min(X, Y) \le w]$$

$$\Rightarrow 1 - F_W(w) = P[\min(X, Y) > w]$$

$$= P[X > w \overset{\text{and}}{,} Y > w]$$

$$= \int_w^\infty \int_w^\infty f_{X,Y}(x, y)\, dx\, dy$$

$$\Rightarrow F_W(w) = 1 - \int_w^\infty \int_w^\infty f_{X,Y}(x, y)\, dx\, dy \quad \#$$

If $\vec{\mathbb{T}}$ and $\vec{\mathbb{G}}$ are linear operators,

then $\vec{\mathbb{T}}\left(\vec{\mathbb{G}}(x,y)\right) = \vec{\mathbb{G}}\left(\vec{\mathbb{T}}(x), \vec{\mathbb{T}}(y)\right)$.

Expectation is linear.

---

(4) $X \overset{d}{=} Y$ means "X and Y are equivalent in distribution", i.e.,
$F_X(x) = F_Y(y)$.

(5) $X = Y \Rightarrow X \overset{P}{=} Y \Rightarrow X \overset{d}{=} Y$
$\Rightarrow X \overset{MS}{=} Y \Rightarrow X \overset{d}{=} Y$