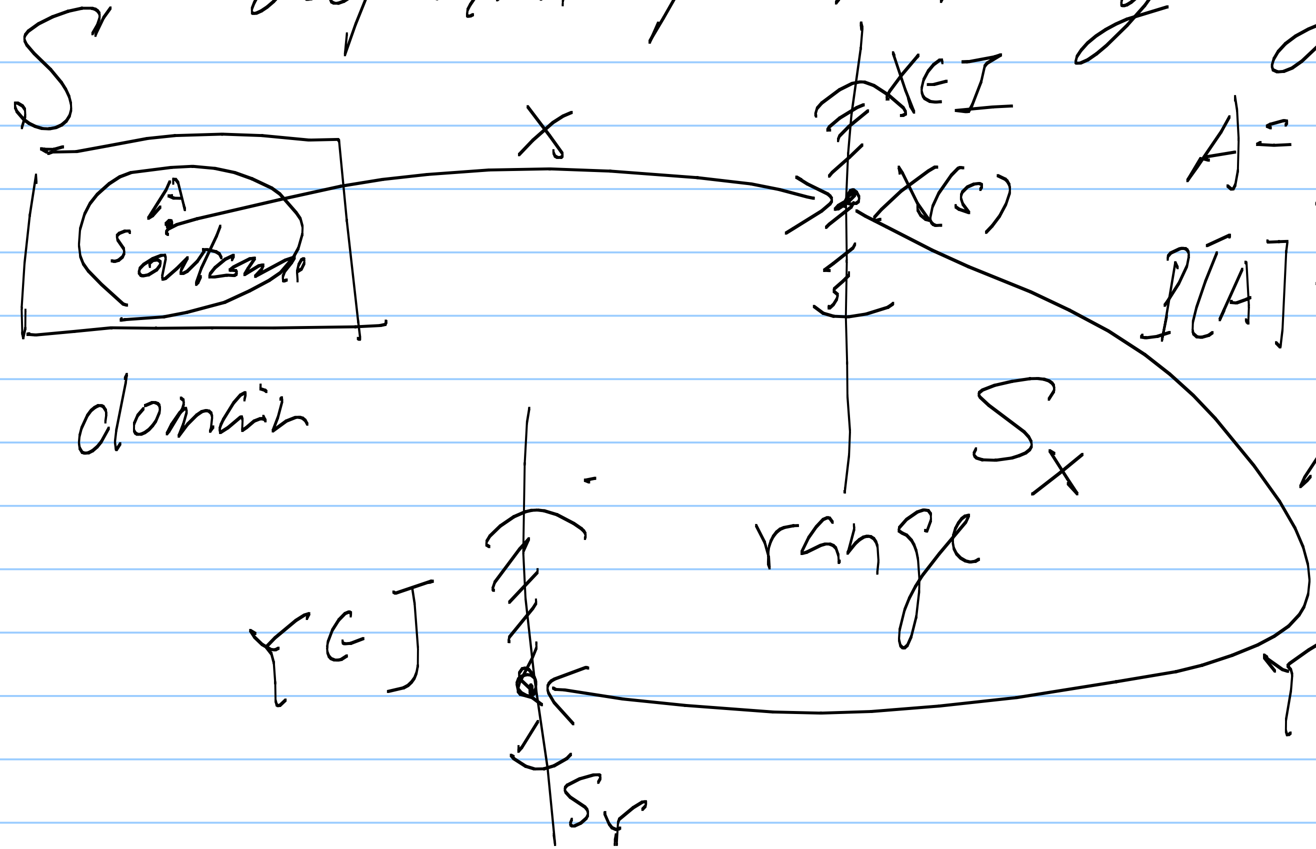
CHAP 4

Recall: One Random Variable is defined pictorially by



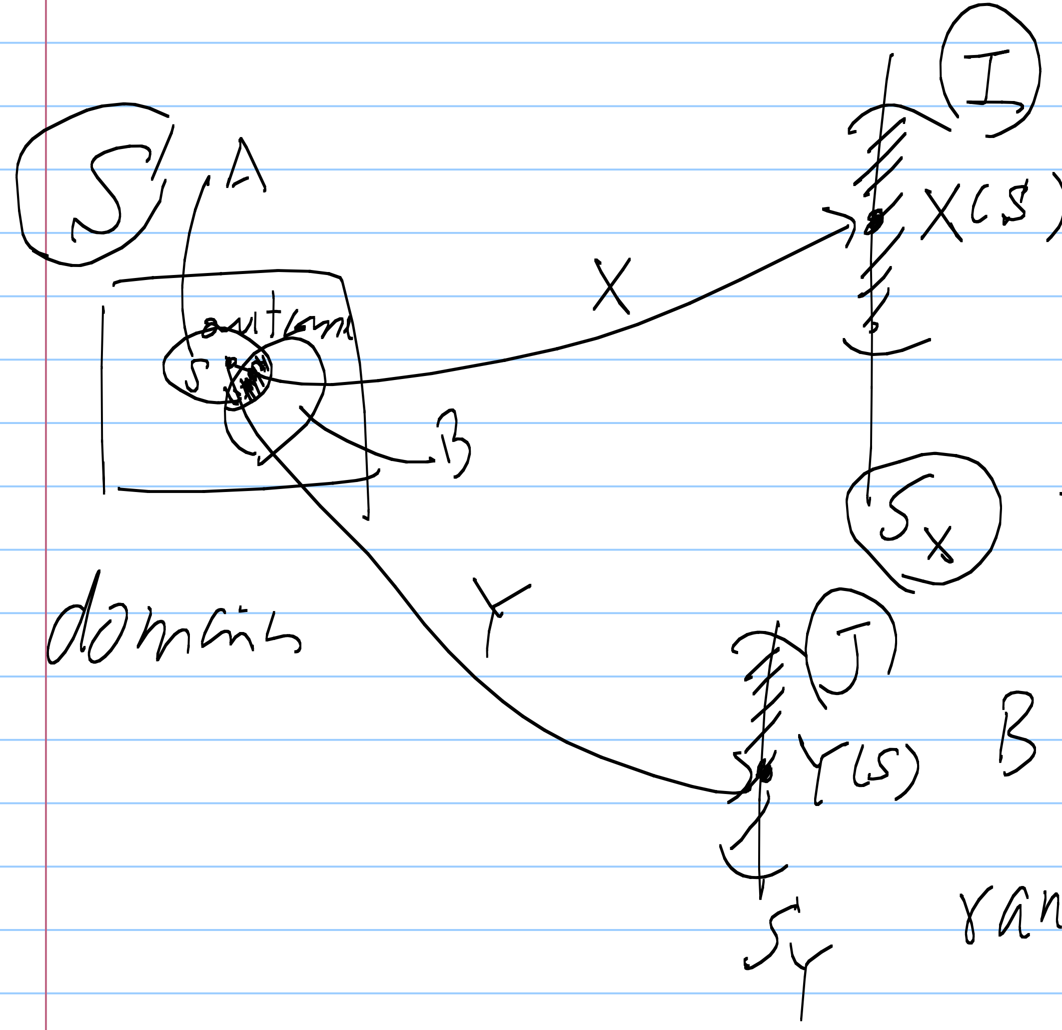
$A = \{X \in I\}$

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

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

$Y = g(X)$

Define Two Random Variables Pictorally:



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

$P[A \cap B]$

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

intersection

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

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

$\equiv S_{X,Y}$

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

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 $y \le y_1$,

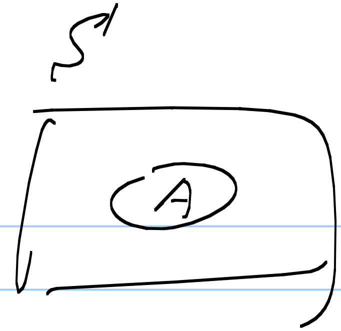
$\Rightarrow F_{X,Y}(x,y) = P[\underbrace{X \le x}_{A}, \underbrace{Y \le y}_{B}] \le P[X \le x_1, Y \le y_1] = F_{X,Y}(x_1, y_1)$

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

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

pf: $F_{X,Y}(\infty,\infty) = P[X \le \infty, Y \le \infty] = P[S \cap S] = P[S] = 1$

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

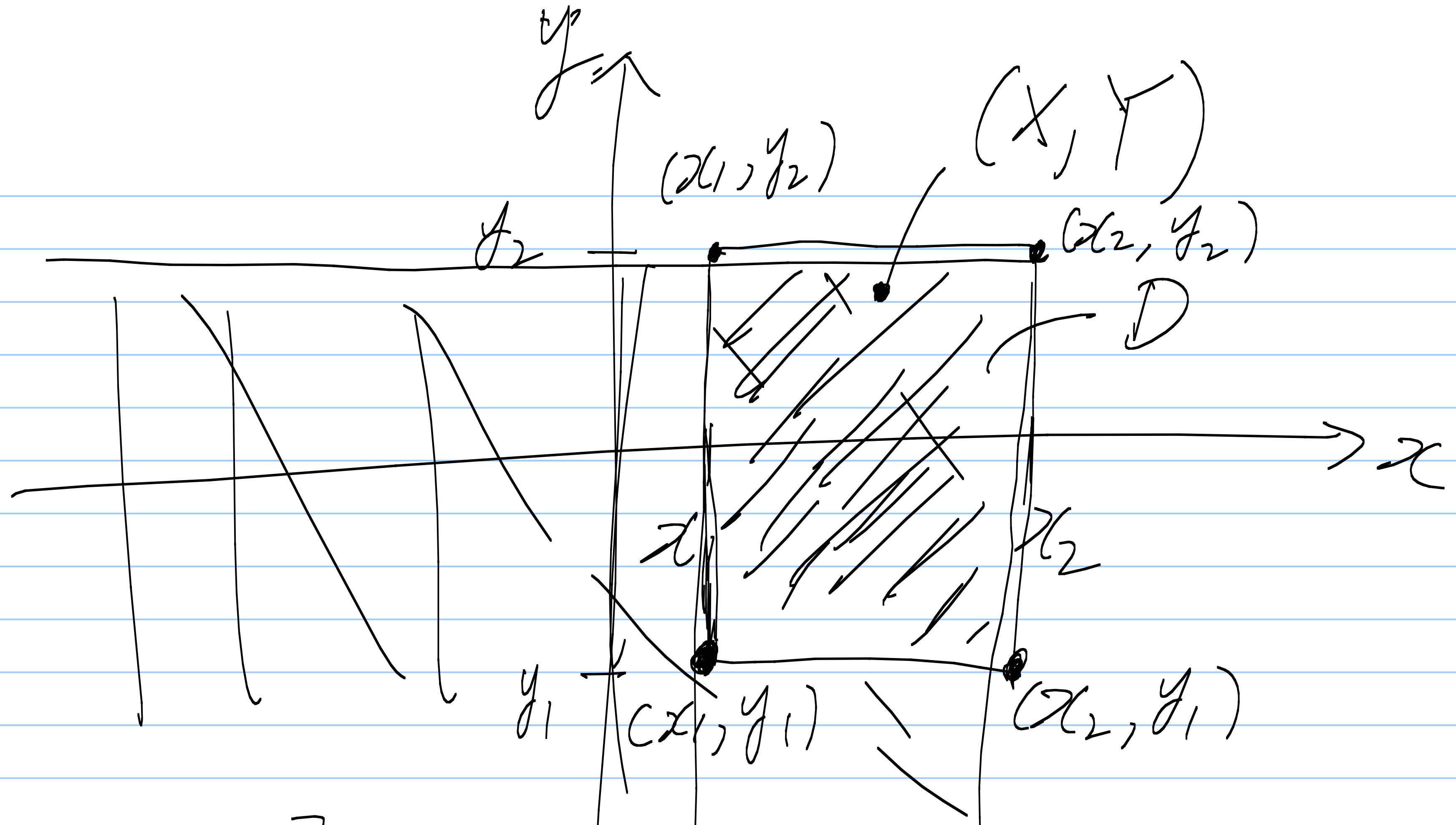


$$P_X(x) \stackrel{\circ}{=} P[\underbrace{X=x}_{A}] = P[A \cap \underbrace{S}_{Y\in S_Y}]$$

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

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

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.



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

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

2-dimensional region

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

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

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

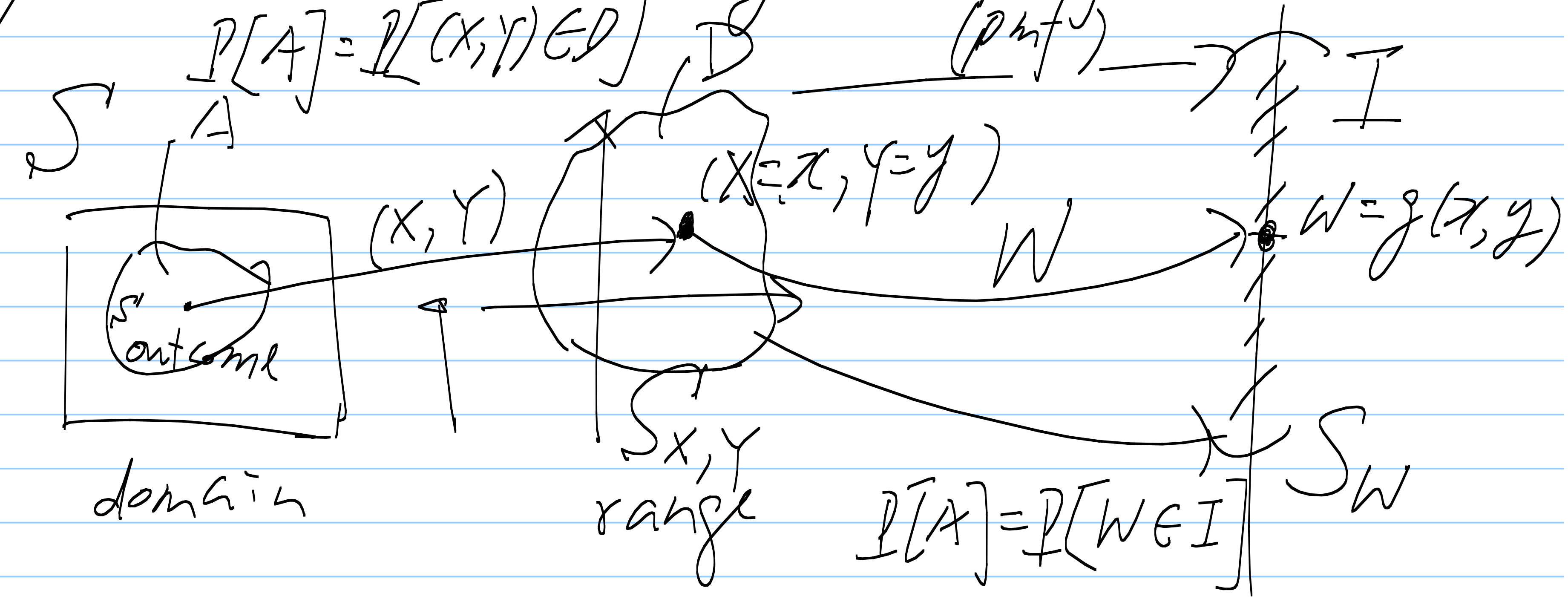
marginal CDF of Y

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

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

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

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

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

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

$$F_W(w) \overset{\Delta}{=} P[W \le w] = P[\max(X, Y) \le w]$$

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

and

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

Q.E.D.

By law of unconscious statistician,

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

(i) If X and Y are discrete,

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

(ii) If X and Y are continuous,

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

$$\sigma^2_{X,Y} \equiv \mathrm{Cov}[X, Y]$$

$$\stackrel{\wedge}{=} E[(X-\mu_X)(Y-\mu_Y)]$$

summation

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

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

If $\mathrm{Cov}[X, Y] = 0$, then

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

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

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

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

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

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

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

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

#

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

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

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

$$\rightarrow \underbrace{f_{X,Y|B}(x,y)\,dx\,dy}\ \approx P[\,x\le X< x+dx,\ y\le Y< y+dy \mid B\,]$$

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

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

or

$B = \{w \le W < w + dw\}$ for any typed random variable W and for $dx \to 0$.

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

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

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

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

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

#

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

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

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

If Y is discrete,

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

$*$ $E[g(X,Y)] = E\big[E[g(X,Y)\mid X]\big]$

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

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

* Recall that two events A and B are independent iff (if and only if) $P[A \cap B] = P[A]\,P[B]$.

* Defn: Two discrete random variables X and Y are called independent iff $P[X=x, Y=y] = P[X=x]\,P[Y=y]$ for all x 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 $dy \to 0$,

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

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

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 $\Leftarrow$ independent

uncorrelated $\not\Rightarrow$ independent in general

uncorrelated $\xRightarrow{\text{bivariate Gaussian}}$ independent

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,

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

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

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

You can also write

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

where $f_Y(y) \sim G(\mu_2, \sigma_2^2)$

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

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

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

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,

$$\mathrm{Var}[Y \mid X = x] = \sigma_2^2 (1 - \rho^2) \leq \sigma_2^2 = \mathrm{Var}[Y]$$

$$\mathrm{Var}[X \mid Y = y] = \sigma_1^2 (1 - \rho^2) \leq \sigma_1^2 = \mathrm{Var}[X]$$

#