① .

Multiple Random Variables

Ch. 4
L.-G.

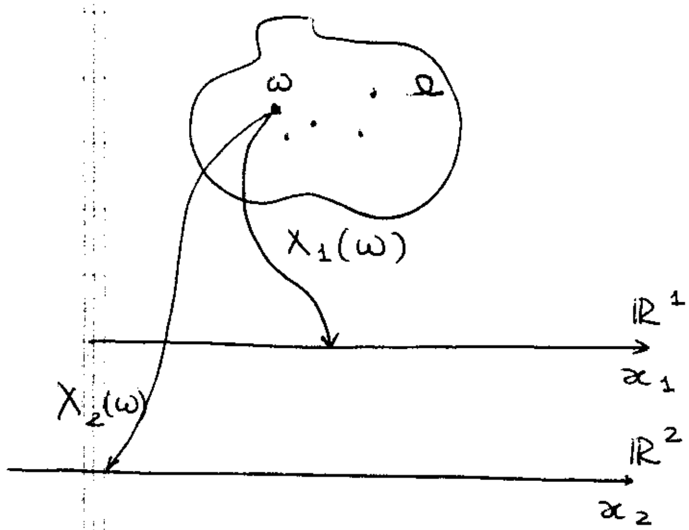
ex. 1: measuring different parameters
(few observable parameters) in
a control system.

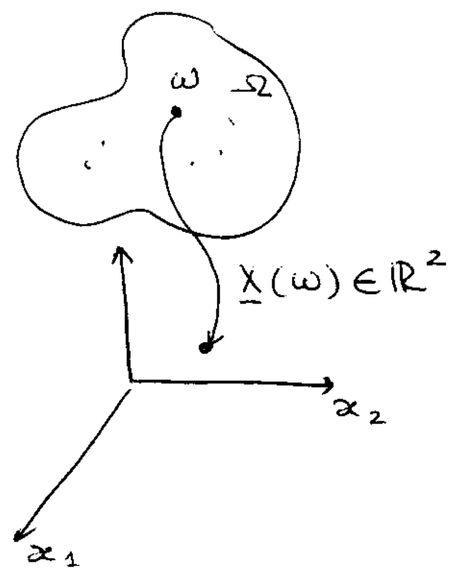
ex. 2: Taking multiple samples
(sequences of trials)

- fundamental notions
  about multiple RVs

- focus on 2 RVs
  (allows geometric interpretation)

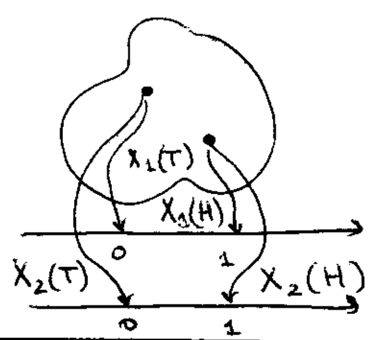
Defn.: A vector R.V. $\underline{X}$ is a function that assigns a vector of real numbers to each outcome $\omega \in \Omega$





ex. 1: Toss two coins. Note outcomes (experiment).

$\Omega = \{ HH, HT, TH, TT \}$



[redacted]

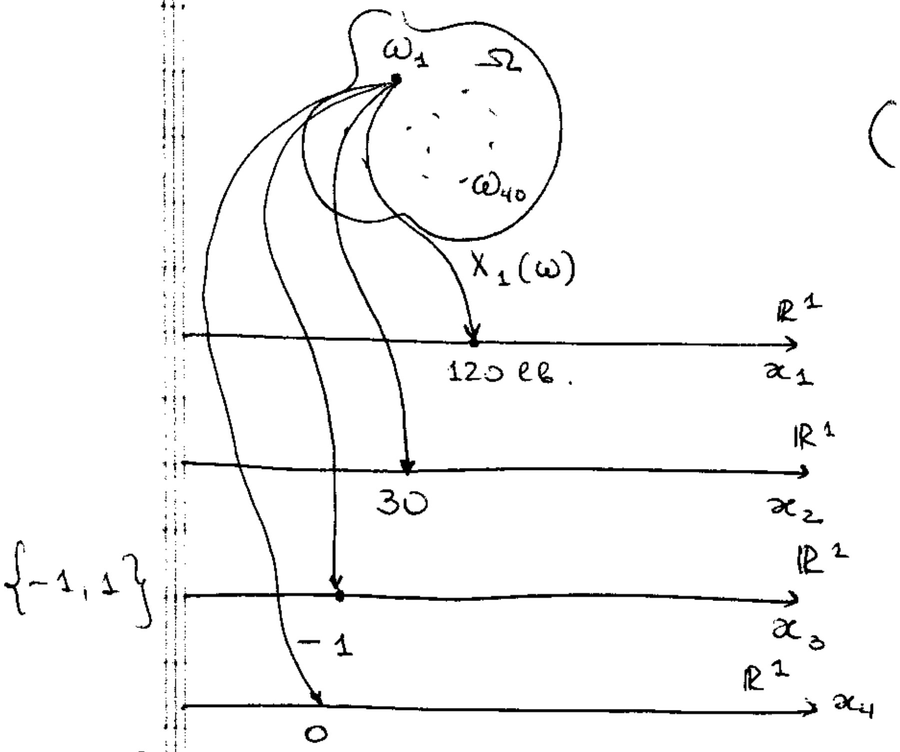
2.

ex. 2: (Group of patients)

Statistics on patients. obesity

40 individuals are observed.

Records : weight, age, gender, etc. (1 sick or 0 healthy)



A triplet :

$$\left(X_1(\omega),\ X_2(\omega),\ X_3(\omega) \right) \quad X_4(\omega)$$

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

(Random Process)

ex. 3: $X(t) = A\cos(\omega t + \varphi)$ RVs

$X(t_1), X(t_2), \ldots, X(t_n)$

$$\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

Events: Each event in $\mathbb{R}^2$ ($\mathbb{R}^n$ in general) is assigned a region in $\mathbb{R}^2$ ($\mathbb{R}^n$ in general).

ex. 1: (two coins)

$A = \{$ Two tosses output one "H" and one "T" $\} = \{HT, TH\}$

$B = \{$ Same outcomes are observed for both tosses $\}$

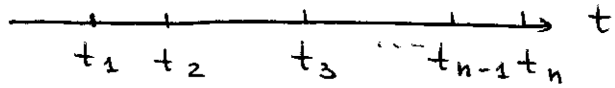
$= \{HH, TT\}$

ex. 2:

$A = \{$ healthy women of age 50 and younger, weight < 140 lb. $\}$

$B = \{$ healthy individuals of age > 30 weight > 200 lb. $\}$

[redacted]



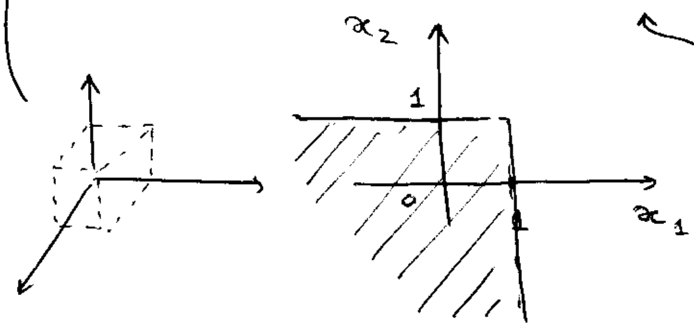
A = { all samples take value below 1 }

B = { Three first samples take values between 1 and 5, the rest drop below 1 }

$\{ X_1 \le 1,\ X_2 \le 1, \ldots, X_n \le 1 \}$

$\{ 1 \le X_1 \le 5,\ 1 \le X_2 \le 5,\ 1 \le X_3 \le 5,\ X_4 \le 1, \ldots, X_n \le 1 \}$

intersection



Note that the events of interest are of <u>product</u> form:

$$\{X_1 \in A_1\} \cap \{X_2 \in A_2\} \cap \ldots \cap \{X_n \in A_n\}$$

$$= \{X_1 \in A_1 \cap X_2 \in A_2 \cap \ldots \cap X_n \in A_n\}$$

$$= \{X_1 \in A_1,\ X_2 \in A_2, \ldots, X_n \in A_n\}$$

Thus A occurs only if each of individual events occurs (that is, they occur jointly).

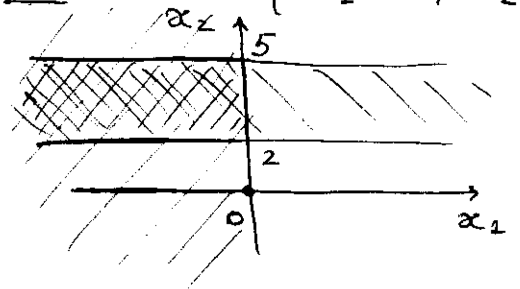
- When modeling a system characterized by a vector R.V. $\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$, a fundamental problem is to specify the probability of product form events.

[redacted]

(4). • For a single R.V. we expressed any event A in terms of operations $(\cup, \cap, /)$ on <u>semi-open intervals</u>.

• In 2-D, all events (product form) can be expressed using <u>semi-infinite rectangles</u>.

ex.: $A = \{X_1 \leq 0,\ X_2 \in [2,5]\}$



To <u>completely specify</u> a vector R.V. $\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ we need to know its <u>joint cumulative distribution function</u>:

$$P_{\underline{X}}(\underline{x}) \triangleq \Pr[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n]$$

$$= \Pr[\omega : X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n]$$

$$= P_{X_1,\ldots,X_n}[x_1, \ldots, x_n]$$

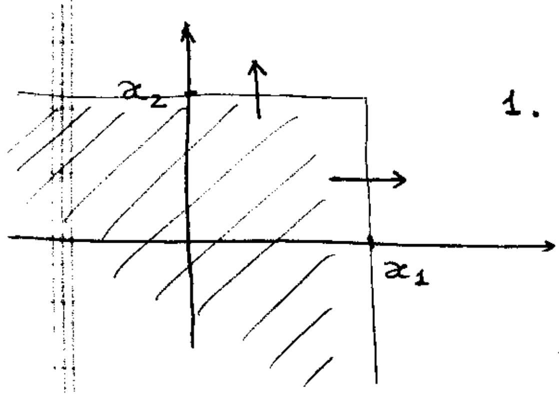
<u>Pairs of R.V.s</u> (2D)

$$\underline{X}(\omega) = \begin{bmatrix} X_1(\omega) \\ X_2(\omega) \end{bmatrix}$$

The c.d.f.

$$P_{\underline{X}}(\underline{x}) \triangleq \Pr[X_1 \leq x_1,\ X_2 \leq x_2]$$

<u>Properties of the cdf</u> (2D):



1. $$\lim_{\substack{x_1 \to +\infty \\ x_2 \to +\infty}} P_{X_1,X_2}(x_1, x_2) = \lim_{\substack{x_1 \to +\infty \\ x_2 \to +\infty}} \Pr[X_1 \leq x_1,\ X_2 \leq x_2]$$

$$= \lim_{\substack{x_1 \to +\infty \\ x_2 \to +\infty}} \Pr[\omega : X_1(\omega) \leq x_1 \cap X_2(\omega) \leq x_2]$$

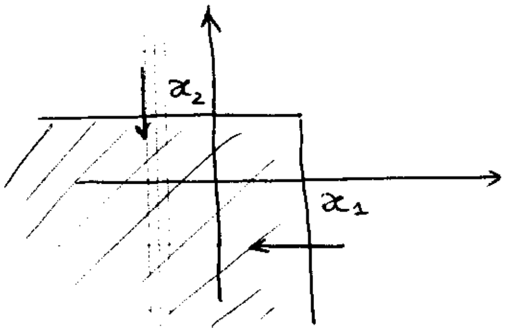
$$= \lim_{\substack{x_1 \to +\infty \\ x_2 \to +\infty}} \Pr[\underbrace{\{\omega : X_1(\omega) \leq x_1\}}_{\Omega} \cap \underbrace{\{\omega : X_2(\omega) \leq x_2\}}_{\Omega}] = 1.$$

[redacted]

2. $\lim\limits_{\substack{x_1 \to -\infty \\ x_2 \to -\infty}} P_{X_1,X_2}(x_1, x_2) = \Pr[\, X_1 \le x_1 ,\ X_2 \le x_2 \,]$

$$= \lim_{\substack{x_1 \to -\infty \\ x_2 \to -\infty}} \Pr\Big[\underbrace{\{\omega : X_1 \le x_1\}}_{\to \emptyset} \cap \underbrace{\{\omega : X_2 \le x_2\}}_{\to \emptyset} \Big]$$

$$= 0$$



3. $\lim\limits_{x_1 \to +\infty} P_{X_1,X_2}(x_1, x_2)$

$$= \lim_{x_1 \to +\infty} \Pr\Big[\underbrace{\{\omega : X_1(\omega) \le x_1\}}_{\to \Omega} \cap \{\omega : X_2 \le x_2\} \Big]$$

$$= \Pr\{ X_2 \le x_2 \} = P_{X_2}(x_2)$$

Similarly, $\lim\limits_{x_2 \to +\infty} P_{X_1,X_2}(x_1, x_2) = P_{X_1}(x_1)$

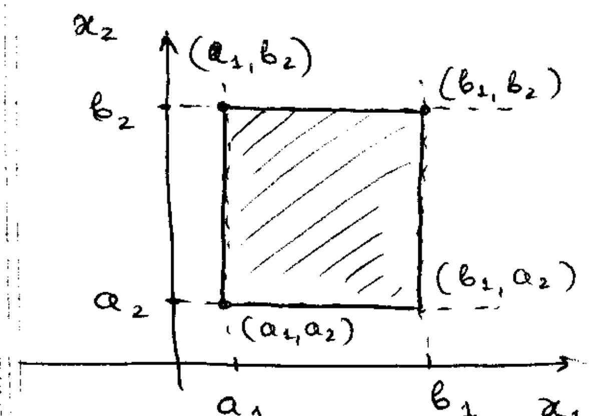
4. $\lim\limits_{x_1 \to -\infty} P_{X_1,X_2}(x_1, x_2) = \lim\limits_{x_1 \to -\infty} \Pr\Big[\underbrace{\{\omega : X_1(\omega) \le x_1\}}_{\emptyset} \cap \{\omega : X_2(\omega) \le x_2\} \Big]$

$$= 0$$

Similarly, $\lim\limits_{x_2 \to -\infty} P_{X_1,X_2}(x_1, x_2) = 0$

Note : Given $P_{X_1,X_2}(x_1, x_2)$, we can find $P_{X_1}(x_1)$ and $P_{X_2}(x_2)$ (marginal distributions)

The converse is not valid.

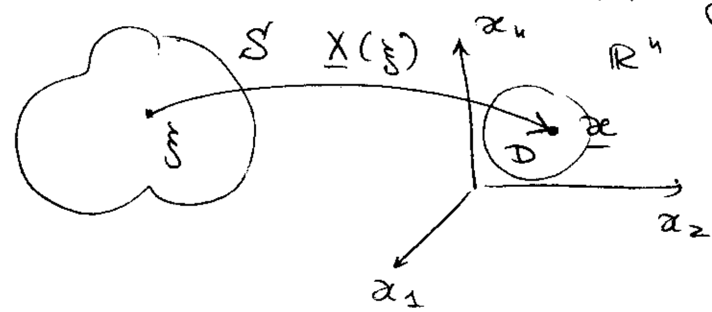
5. 

$$\Pr[a_1 < X_1 \le b_1 ,\ a_2 < X_2 \le b_2]$$
$$= P_{X_1,X_2}(b_1, b_2) - P_{X_1,X_2}(b_1, a_2) - P_{X_1,X_2}(a_1, b_2) + P_{X_1,X_2}(a_1, a_2)$$

[redacted]

Summary: [10/23/2008]

- Measurements are taken simultaneously at different locations in a complex stochastic system.
- They have to be processed jointly
- A vector R.V. $\underline{X}$ is a mapping



- Events are subsets of $\mathbb{R}^n$,

$$D \subset \mathbb{R}^n$$

- To assign probability to events, we focus on product type events.

If

$$A = \{X_1 \in A_1, X_2 \in A_2, \ldots\ldots, X_n \in A_n\},$$

"$\cap$"

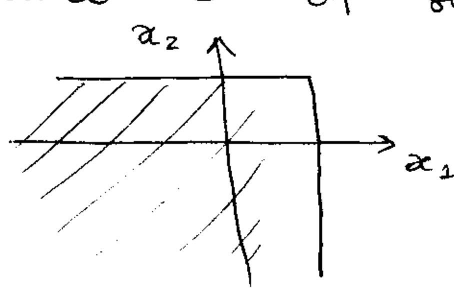
then A is of product type

ex. (nonproduct type)

$$A = \{\max(X_1, \ldots\ldots, X_n) \leq \alpha\}$$

- Product type events can be expressed as $\cup$, $\cap$ and "c" of semiopen rectangles.

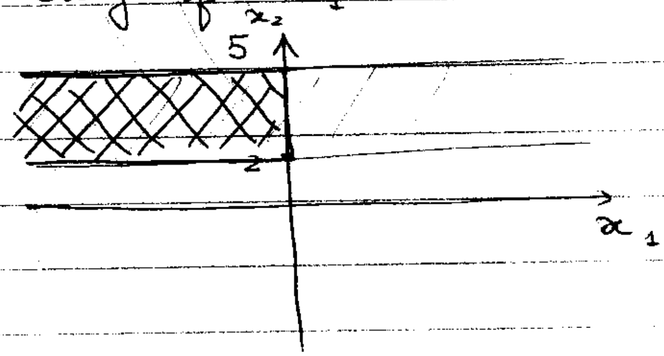
"Building block"



Then

$$A = \{ X_1 \le 0 ,\ X_2 \in [2,5] \}$$

occurs only if $X_1 \le 0$ and $X_2 \in [2,5]$



- We define the joint cdf as

$$F_{X_1, X_2}(x_1, x_2) \triangleq P[X_1 \le x_1 ,\ X_2 \le x_2]$$

probability of semiopen rectangle

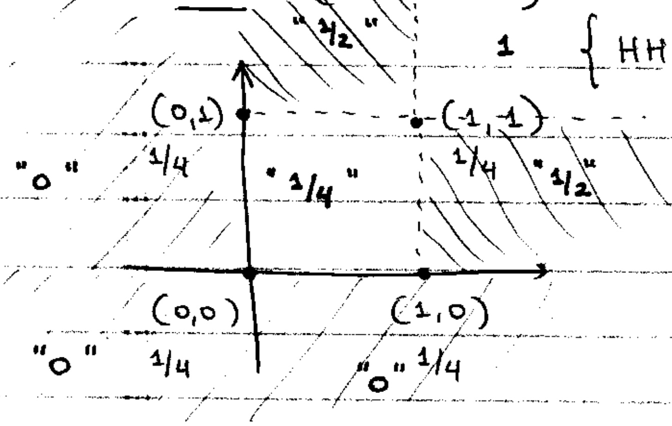
- Properties of cdf ....

<u>ex</u>: $P[A] = P[X_1 \le 0 ,\ X_2 \in [2,5]]$

$$= F_{X_1,X_2}(0,5) - F_{X_1,X_2}(0,2)$$

<u>ex</u>. (See the next page)

⑥ <u>ex.</u> (2 coins)



$\{HH, HT, TH, TT\}$

$x_1 \quad , \quad x_2 < 0$

$x_1 < 0 \quad , \quad x_2$

If $x_1 < 0$, $x_2 < 0$, then

$P_{X_1, X_2}(x_1, x_2) = 0$

If $0 \le x_1 < 1$, $0 \le x_2 < 1$

$P_{X_1, X_2}(x_1, x_2) = 1/4$

If $0 \le x_1 < 1$, $x_2 \ge 1$, then $P_{X_1, X_2}(x_1, x_2) = 1/2$

If $0 \le x_2 < 1$, $x_1 \ge 1$, then $P_{X_1, X_2}(x_1, x_2) = 1/2$

If $x_1 \ge 1$, $x_2 \ge 1$, then $P_{X_1, X_2}(x_1, x_2) = 1$

<u>Theorem:</u> A function $G_{X_1, X_2}(x_1, x_2)$ is a valid jdf iff

(1) it is nondecreasing in x_1 and x_2 from 0 to 1,

(2) for all order choices a_1, a_2, b_1, b_2

s.t. $b_1 \ge a_1$, $b_2 \ge a_2$

$Pr[a_1 < X_1 \le b_1, \; a_2 < X_2 \le b_2] \ge 0$

<u>Joint probability density function</u>:

$$f_{X_1, X_2}(x_1, x_2) \triangleq \frac{\partial^2}{\partial x_1 \partial x_2} P_{X_1, X_2}(x_1, x_2),$$

if the derivative of jdf exists.

Then the jdf can be expressed as the following integral of the pdf:

$$P_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1, X_2}(\alpha, \beta) \, d\alpha \, d\beta$$

[redacted]

Properties of joint pdf :

- $f_{X_1,X_2}(x_1,x_2) \geq 0$
- marginal pdfs : $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$

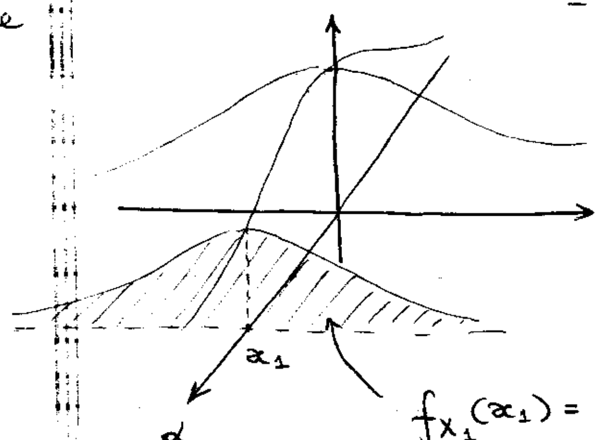
Take a look at jdf :

$$P_{X_1,X_2}(x_1,\infty) = P_{X_1}(x_1) = \int_{-\infty}^{x_1}\left[\int_{-\infty}^{+\infty} f_{X_1,X_2}(\alpha,\beta)\,d\beta\right]d\alpha$$

Differentiate $P_{X_1,X_2}(x_1,\infty)$ w. r. to x_1 :

Leibnitz rule

$$f_X(x_1) = \int_{-\infty}^{+\infty} f_{X_1,X_2}(x_1,\beta)\,d\beta$$



$$f_{X_1}(x_1) = \underbrace{\int_{-\infty}^{+\infty} f_{X_1,X_2}(x_1,\beta)\,d\beta}_{\text{area}}$$

Exercise 3.1 :

X, Y, Z ~ independent

(a) $\Pr(|X|<5,\ Y>2,\ Z^2 \geq 2)$ $\{z \geq \sqrt{2} \cup z \leq -\sqrt{2}\}$

$= \Pr[|X|<5]\cdot\Pr[Y>2]\cdot\Pr[Z^2\geq 2]$

$= [P_X(5^-) - P_X(-5)]\cdot[1 - P_Y(2)]$

$\times [P_Z(-\sqrt{2}) + 1 - P_Z(\sqrt{2}^-)]$

(b) $\Pr(X>5,\ Y<0,\ Z=1)$

$= [1 - P_X(5)]\cdot P_Y(0^-)\cdot[P_Z(1) - P_Z(1^-)]$

(c) $\Pr[\min(X,Y,Z) > 2] = \Pr[X>2,\ Y>2,\ Z>2]$

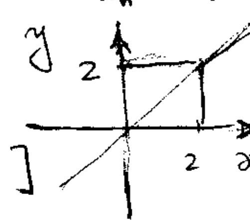
$\underset{\text{independence}}{=} \Pr[X>2]\Pr[Y>2]\Pr[Z>2]$

$= [1 - P_X(2)][1 - P_Y(2)][1 - P_Z(2)]$

Aside: n random variables $X_1, \ldots, X_n$ are independent if

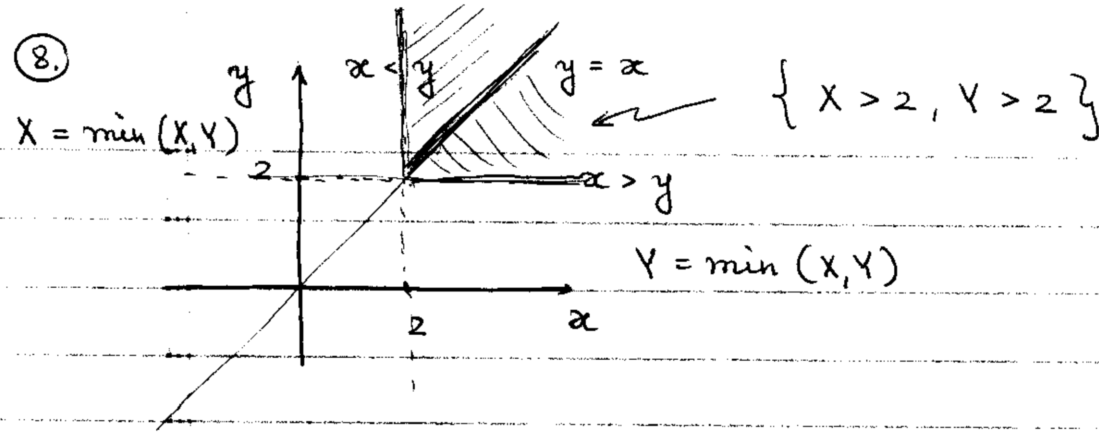
$P_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$

$= P_{X_1}(x_1)\times\ldots\times P_{X_n}(x_n)$



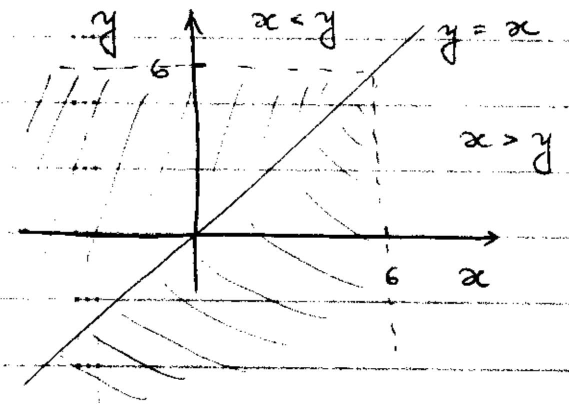
[redacted]

(8.)



(d). $\Pr[\max(X,Y,Z) < 6] = \Pr[X < 6, Y < 6, Z < 6]$

$= P_X(6^-)\, P_Y(6^-)\, P_Z(6^-)$



Exercise 3.2 : $X \perp\!\!\!\perp Y \quad U([0,1])$

(a) $\Pr\left\{X^2 < \frac{1}{2}, |Y-1| < \frac{1}{2}\right\}$

$$= \Pr\left\{-\frac{1}{\sqrt{2}} < X < \frac{1}{\sqrt{2}}\right\} \cdot \Pr\left\{\frac{1}{2} < Y < \frac{3}{2}\right\}$$

$$= \int_0^{1/\sqrt{2}} dx \cdot \int_{1/2}^{1} dy = \frac{1}{\sqrt{2}} \cdot \frac{1}{2} = \frac{1}{2\sqrt{2}}$$

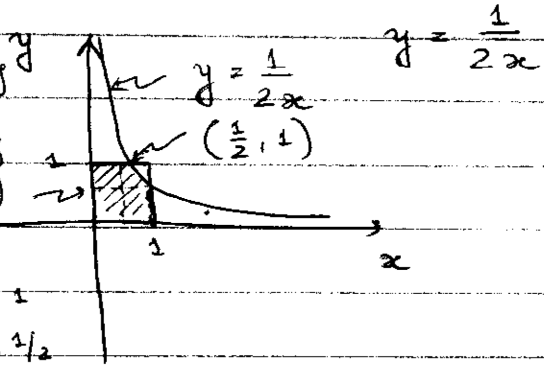
(b). $\Pr\left\{\frac{1}{2}X < 1, Y > 0\right\} = \Pr\{X < 2\} \cdot \Pr\{Y > 0\}$

$= 1 \cdot 1 = 1$

not a product ~~type~~

(c) $\Pr\left\{XY < \frac{1}{2}\right\} = \iint f_{X,Y}\, dx\, dy$

$\left\{(x,y) : xy < \frac{1}{2}\right\}$



$y = \frac{1}{2x}$

$$= \frac{1}{2} + \int_{1/2}^{1} \int_0^{1/2x} dy\, dx$$

$x \quad y$

$$= \frac{1}{2} + \int_{1/2}^{1} \frac{1}{2x}\, dx = \frac{1}{2} + \frac{1}{2} \ln x \Big|_{1/2}^{1}$$

$$= \frac{1}{2}(1 + \ln 2)$$

[redacted]

(9). (d) $\Pr\{\min(X,Y) > \frac{1}{3}\} = \Pr\{X > \frac{1}{3},\ Y > \frac{1}{3}\}$

$$= \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$$

X Exercise 3.3: (next time)

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x} \cdot e^{-y} & , 0 \le y \le x < \infty \\ 0 & , \text{otherwise} \end{cases}$$

Find: the pdf of $Z = X + Y$

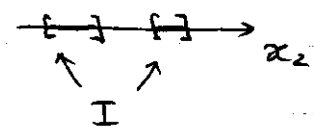
Conditional distribution and pdf

$$P(A \mid B) \triangleq \frac{P(A,B)}{P(B)} \quad P(B) \neq 0$$

A and B are two events

Suppose $A = \{\omega : X_1(\omega) \le x_1\}$

$B = \{\omega : X_2(\omega) \in I\}$



(1) The conditional probability of A given B:

$\Pr[X_1 \le x_1 \mid X_2 \in I]$

$$\triangleq F_{X_1}(x_1 \mid X_2 \in I)$$

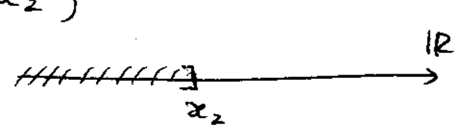
$$= \frac{\Pr[X_1 \le x_1, X_2 \in I]}{\Pr[X_2 \in I]} = \frac{\int_{-\infty}^{x_1} \int_{\beta : \beta \in I} f_{X_1,X_2}(\alpha,\beta)\, d\alpha\, d\beta}{\int_{\beta : \beta \in I} f_{X_2}(\beta)\, d\beta}$$

if $f_{X_2}(x_2) \neq 0$

The pdf of X_1, given $x_2 \in I$:

$$f_{X_1}(x_1 \mid X_2 \in I) \triangleq \frac{d}{dx_1} F_{X_1}(x_1 \mid X_2 \in I)$$

(2) Suppose that $I = (-\infty, x_2)$



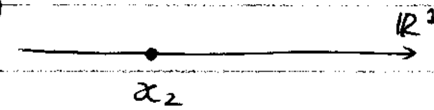
$\Pr[X_1 \le x_1 \mid X_2 \le x_2]$

[redacted]

$$\triangleq F_{X_1|X_2}(x_1|x_2) = \frac{\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1,X_2}(\alpha,\beta)\,d\alpha\,d\beta}{\int_{-\infty}^{x_2} f_{X_2}(\beta)\,d\beta}$$

$$= \frac{F_{X_1,X_2}(x_1,x_2)}{F_{X_2}(x_2)}$$

(3). Suppose that $I = \{x_2\}$



If X is a continuous RV, then

$\Pr[X = x] = 0$.

Consider a very short interval:

$$I = (x_2 - \Delta, x_2]$$

The conditional cdf of X_1 given that $X_2 \in (x_2-\Delta, x_2]$

$$F_{X_1}(x_1 \mid X_2 \in (x_2-\Delta, x_2])$$

$$= \frac{\Pr[X_1 \le x_1, X_2 \in (x_2-\Delta, x_2]]}{\Pr[X_2 \in (x_2-\Delta, x_2]]}$$

If Δ is small, the numerator $\simeq \frac{\partial}{\partial x_2} F_{X_1,X_2}(x_1,x_2) \times \Delta$

$$= \frac{F_{X_1,X_2}(x_1,x_2) - F_{X_1,X_2}(x_1,x_2-\Delta)}{F_{X_2}(x_2) - F_{X_2}(x_2-\Delta)} \cdot \frac{\Delta}{\Delta}$$

$$\lim_{\Delta\to 0}(\quad) = \frac{\frac{\partial}{\partial x_2} F_{X_1,X_2}(x_1,x_2)}{\frac{\partial}{\partial x_2} F_{X_2}(x_2)}$$

The conditional pdf of X_1 given $X_2 = x_2$:

$$\boxed{f_{X_1|X_2}(x_1|x_2) = \frac{\partial}{\partial x_1} F_{X_1}(x_1 \mid X_2 = x_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}}$$

[redacted]

Together with cdf, we use pmf and pdf to completely characterize a descrete or continuous RV.

Let $\underbrace{X}_{\text{values } a_k}$ and $\underbrace{Y}_{\text{values } b_k}$ be a pair of discrete RVs.

The event $A = \{X = a_k,\ Y = b_\ell\}$ is of product form.

$$P[X = a_k,\ Y = b_\ell] = p_{X,Y}(a_k, b_\ell) \quad \sim \text{ joint pmf}$$

In terms of joint cdf

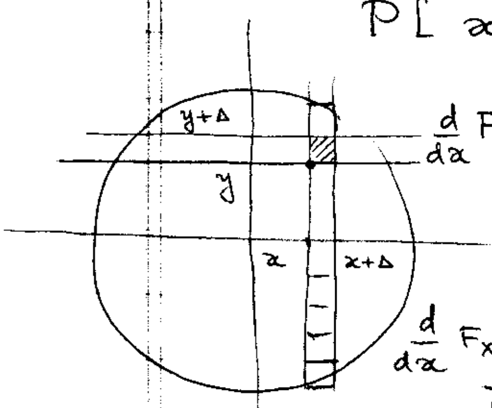
$$p_{X,Y}(a_k, b_\ell) = F_{X,Y}(a_k, b_\ell) - P[X < a_k,\ Y < b_\ell]$$

For continuous RVs and non-product type of events, the sets describing events can be partitioned into rectangles of a small width Δ.

Rectangles

$$P[x \le X \le x + \Delta_x,\ y \le Y \le y + \Delta_y]$$

can be expressed in terms of jcdf.



$$\overbrace{F_{X,Y}(x+\Delta_x, y+\Delta_y) - F_{X,Y}(x, y+\Delta_y)}^{\frac{d}{dx} F_{X,Y}(x, y+dy)}$$
$$\underbrace{- F(x+\Delta_x, y) + F_{X,Y}(x, y)}_{\frac{d}{dx} F_{X,Y}(x,y)}$$
$$\simeq f_{X,Y}(x,y) \cdot \Delta_x \cdot \Delta_y$$

In the limit: $\displaystyle\iint_{x^2+y^2 \le r^2} f_{X,Y}(x,y)\, dx\, dy$

$\frac{d}{dx} F_{X,Y}(x, y+dy)$

$- \frac{d}{dx} F_{X,Y}(x, y)$

Ch. 4: Multiple RVs

Independence

← (prior to conditional pdfs and cdfs)

Defn Two RVs, X and Y, are statistically independent, if the events $\{X \in A\}$ and $\{Y \in B\}$ are statistically independent, that is,

$$\Pr\{X \in A, Y \in B\} = \Pr\{X \in A\} \cdot \Pr\{Y \in B\}.$$

Applying to the events

$$A = (-\infty, x] \quad , \quad B = (-\infty, y],$$

Joint cdf. $F_{X,Y}(x,y) = F_X(x) F_Y(y)$

and

Joint p.d.f. $f_{X,Y}(x,y) = f_X(x) f_Y(y)$

X and Y jointly continuous

marginal pdfs.

Joint p.m.f. (X and Y are discrete)

Events $A = \{X = a_j\}$

$B = \{Y = b_k\}$

$$P_{X,Y}(a_j, b_k) = P_X(a_j) \cdot P_Y(b_k)$$

product of marginal pmf's

Note: If $Z = h(X)$ and ← fuct. of X only

$W = g(Y)$ ← fuct. of Y only

and $X \perp\!\!\!\perp Y$

$\tilde{A}$ and A are equivalent

$\tilde{B}$ and B are equivalent

Then

$$P[Z \in A, W \in B] = P[X \in \tilde{A}, Y \in \tilde{B}]$$

X and Y are independent

$$= P[X \in \tilde{A}] \cdot P[Y \in \tilde{B}] = P[Z \in A] \cdot P[W \in B]$$

Z and W are also independent.

Conditional Expectation

Defn. The conditional expectation of Y given $X = x$ is defined as

$$E[Y \mid X = x] \triangleq \int_{-\infty}^{+\infty} y \cdot f_{Y|X}(y \mid x)\, dy.$$

In many references conditional expectation is addressed as a function of x, that is, $g(x)$. Since x is a value that RV X takes, we may be interested in defining $g(X)$ and find $E[g(X)]$:

$$E[g(X)] = E[\underbrace{E[Y \mid X]}_{\text{unconditional expectation}}]$$

$$E[E[Y \mid X]] = \int_{-\infty}^{+\infty} E[Y \mid x]\, f_X(x)\, dx$$

$$= \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} y\, f_{Y|X}(y \mid x)\, dy \right\} f_X(x)\, dx$$

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

$$= \int_{-\infty}^{+\infty} y \left\{ \underbrace{\int_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dx}_{f_Y(y)} \right\} dy = E[Y]$$

Conclude, therefore,

$$E[Y] = E[E[Y|X]]$$

Procedure is known as "iterated expectations"

Expectation of Any funct. of Y

$$E[h(Y)] = E[E[h(Y)|X]]$$

ex. (Information Theory)

Entropy $H(X)$ is known a measure of uncertainty of a R.V. X (discrete).
(Indicator of how much we can compress data without losses).

$$H(X) \triangleq -E[\log p(X)]$$

pmf

uncertain R.V.
~ uniform

certain RV
~ narrow pdfs

Conditional entropy characterizes communication channel

$$X \longrightarrow \boxed{\text{channel}} \longrightarrow Y$$

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

$$H(Y|X) \triangleq - E[E[\log p_{Y|X}(Y|X)]]$$

$$= - \sum_x E[\log p_{Y|X}(Y|x)] \cdot p_X(x)$$

$$= - \sum_x \sum_y \log p_{Y|X}(y|x) \cdot p_{Y|X}(y|x) p_X(x)$$

Conditional Expectation

Defn.: The conditional expectation of Y given $X = x$ is defined as

$$E[Y \mid X = x] = \int_{-\infty}^{+\infty} y \cdot f_{Y|X}(y \mid x)\, dy$$

To obtain unconditional expectation of the R.V. Y,

$$\int_{-\infty}^{+\infty} dx\, f_X(x) \cdot E[Y \mid X = x] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \cdot f_{Y|X}(y \mid x)\, dy \cdot f_X(x)\, dx$$

$$= \int_{-\infty}^{+\infty} y \cdot \Big[\underbrace{\int_{-\infty}^{+\infty} f_{X,Y}(y, x)\, dx}_{f_Y(y) \,\leftarrow\, \text{marginal}} \Big]\, dy$$

$$= \int_{-\infty}^{+\infty} y\, f_Y(y)\, dy$$

$$= E\big[E[Y \mid X] \big]$$

Conclusion: To obtain unconditional mean from a conditional mean, multiply by $f_X(x)$ and integrate over x.

If we have $\xrightarrow{Y} \boxed{g(\cdot)} \rightarrow g(Y)$

$\uparrow$ deterministic, real

Then $$E[g(Y) \mid X] = \int_{-\infty}^{+\infty} g(y) \cdot f_{Y|X}(y \mid x)\, dy$$

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

$\uparrow$ iterated expectation

[redacted]

(12) ex.: (Info. Theory)

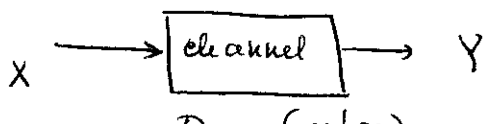
$H(X)$ is a measure of uncertainty of the RV X
(indicator how much we can compress the data without losses)

$$H(X) \triangleq -E[\log p(X)] \leftarrow \text{entropy}$$

uncertain RV. ~ uniform

certain R.V. ~ narrow pdfs.

Communication Channel



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

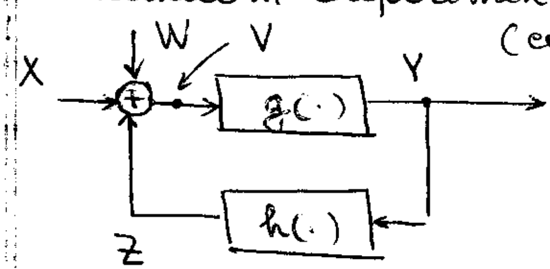
$$H(Y|X) \triangleq -E[\log f_{Y|X}]$$

$$= \iint_{-\infty}^{+\infty} \log \frac{1}{f_{Y|X}(y|x)} \cdot f_{Y|X}(y|x) \cdot f_X(x)\, dx\, dy$$

Functions of RVs

- Random Experiment

ex.



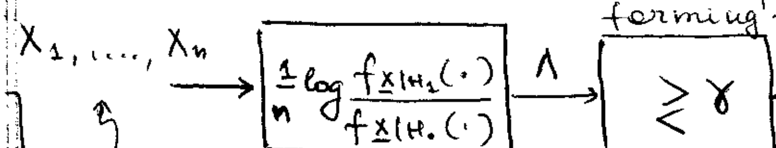
(control system)

RV RV(noise)

$$V = X + W + h(g(V))$$

ex. (Binary receiver)

Given observations that can be related to one of two sources of data and a model for forming the data.



Z ~ binary R.V.

Each source = a hypothesis

The task is to design an optimal detector.

$$\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

observed data with two possible conditional pdfs $f_{\underline{X}|H_0}(\underline{x})$ and $f_{\underline{X}|H_1}(\underline{x})$

[redacted]

(13.)

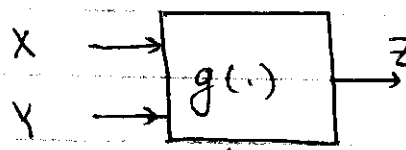
A single function of two RVs

Given two RVs, X and Y, and a function $g(X, Y)$,

(ex. evaluate their empirical mean and variance, i.e., $M = \frac{X+Y}{2}$, $V = \frac{(X-M)^2+(Y-M)^2}{2}$)

consider the RV Z:

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



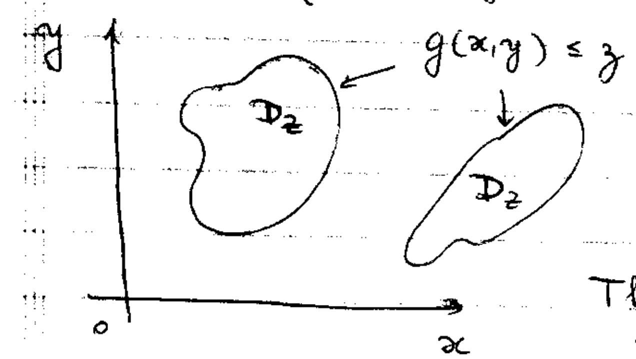
$f_{X,Y}$

deterministic, real valued

<u>Find</u>: cdf and pdf of Z, provided $f_{X,Y}(x, y)$ and $g(\cdot)$ are known.

<u>Steps</u>:
1. Find equivalent events
2. Evaluate cdf and take the derivative w.r. to z.

In general, let D_z be the region in (x, y)-plain that satisfies $D_z = \{ (x,y) : g(x,y) \le z \}$



$$\{ Z \le z \} = \{ g(X, Y) \le z \}$$
$$= \{ (X, Y) \in D_z \}$$

Then

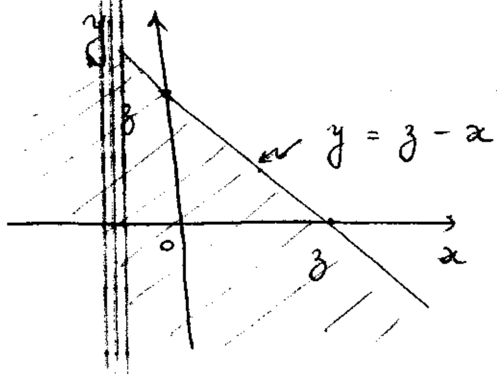
$$P_Z(z) = \Pr[Z \le z]$$
$$= \Pr\{ (X, Y) \in D_z \}$$
$$= \iint_{D_z} f_{X,Y}(x, y)\, dx\, dy$$

Thus, to find $P_Z(z)$, it is sufficient to find the region D_z (an equivalent event) for every z and evaluate the integral.

[redacted]

ex. 1 : Let $Z = X + Y$

X, Y have a joint density $f_{X,Y}(x,y)$



$$\Pr[Z \le z] = \Pr[X + Y \le z]$$

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

"x" "y"

To find the pdf of Z, $f_Z(z)$, differentiate w.r. to z:

$$f_Z(z) = \frac{d\,P_Z(z)}{dz} = \int_{-\infty}^{+\infty} f_{X,Y}(x, z-x)\,dx \qquad (1)$$

Independence and Convolution:

If RVs are independent, then

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

Substituting $f_{X,Y}(x,y)$ into (1), we obtain

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)\cdot f_Y(z-x)\,dx \qquad (2)$$

convolution integral

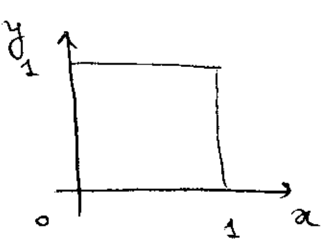
Thus, $Z = X + Y$, $X \perp\!\!\!\perp Y$, the pdf of Z,

$$f_Z(z) = f_X(z) * f_Y(z)$$

ex. 1 : (cntnd)

Let $X \sim \mathcal{U}[0,1]$

$Y \sim \mathcal{U}[0,1]$

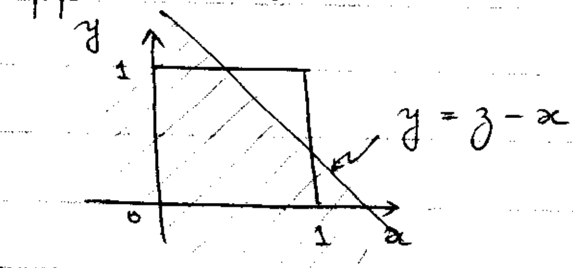


Let $X \perp\!\!\!\perp Y$

independent

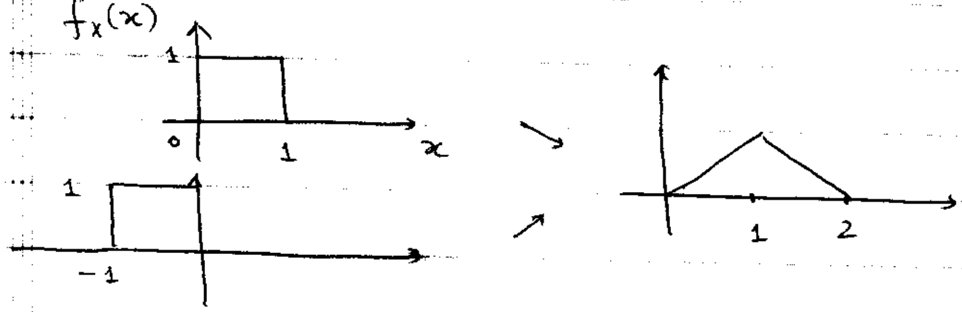
(15). The density of Z can be evaluated using 2 approaches.

(a).

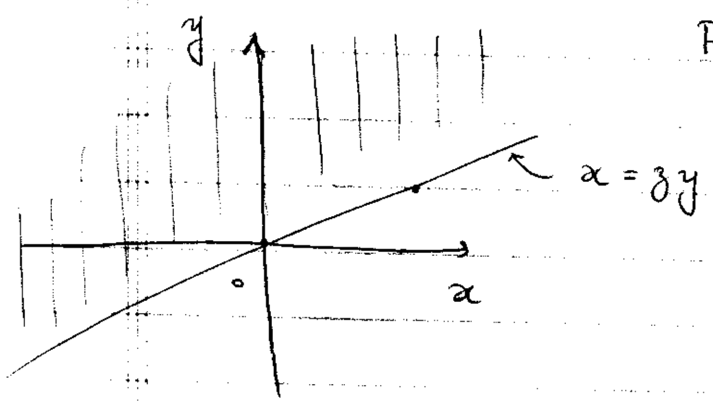


Integrate the region below $x + y \le z$ and differentiate.

(b). Convolve two densities:



ex. 2 : Let $Z = X/Y$



$$Pr[Z \le z] = Pr[X/Y \le z]$$

$x = zy$

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

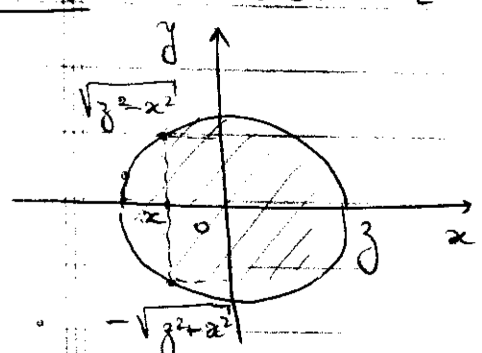
"x" "y"

$(z > 0)$

And the density of Z is

$$f_Z(z) = \frac{1}{z^2}\int_{-\infty}^{+\infty} f_{X,Y}\left(x, \frac{x}{z}\right) \cancel{x} \cdot dx$$

ex. 3 : Let $Z = \sqrt{X^2 + Y^2}$



$$P_Z(z) = Pr[\sqrt{X^2+Y^2} \le z] = Pr[X^2+Y^2 \le z^2]$$

$$= \int_{-z}^{z}\int_{-\sqrt{z^2-x^2}}^{\sqrt{z^2+x^2}} f_{X,Y}(x,y)\, dy\, dx$$

[redacted]