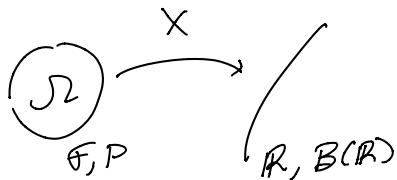


Lecture 8: Measurable functions, Random variables and distribution functions

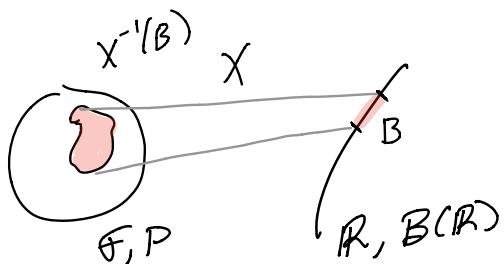
In this lecture we will start by developing the measure theoretic notion of measurable functions & define random variables X as measurable functions $X: \Omega \rightarrow \mathbb{R}$ where (Ω, \mathcal{F}, P) is a prob. space.



Measurability of X is required since we want $P(X \in B)$ to be defined where $B \in \mathcal{B}(\mathbb{R})$ and $\{X \in B\} = \{w \in \Omega : X(w) \in B\}$

$$=: X^{-1}(B)$$

\curvearrowleft
pre-image of B under X .



Measurable functions

Let $(\Omega_1, \mathcal{F}_1)$ & $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces & $f: \Omega_1 \rightarrow \Omega_2$.

Def: f is measurable between \mathcal{F}_1 & \mathcal{F}_2 (written $f @ \mathcal{F}_1/\mathcal{F}_2$ for short) iff

$$f^{-1}(A) \in \mathcal{F}_1, \forall A \in \mathcal{F}_2. \quad (*)$$

Note: It will sometimes be convenient to write $f @ \mathcal{F}_1/\mathcal{F}_2$ when f satisfies (*) even when \mathcal{F}_1 or \mathcal{F}_2 are not σ -fields ... just collections of sets.

A few basic facts about $f^{-1}(A)$

(1) $f^{-1}(\Omega_2) = \Omega_1$ since f maps into Ω_2

(2) $f^{-1}(\emptyset) = \emptyset$

(3) $f^{-1}(A^c) = (f^{-1}(A))^c$
since $w \in f^{-1}(A^c) \Leftrightarrow f(w) \in A^c$
 $\Leftrightarrow f(w) \notin A$
 $\Leftrightarrow w \notin f^{-1}(A)$

(4) $f^{-1}(\bigcup_p A_p) = \bigcup_p f^{-1}(A_p)$ even A_p 's are not disjoint
since $w \in f^{-1}(\bigcup_p A_p) \Leftrightarrow f(w) \in A_p$ some p
 $\Leftrightarrow w \in f^{-1}(A_p)$ some p
 $\Leftrightarrow w \in \bigcup_p f^{-1}(A_p)$

Thm (Generators are enough)

If $\Omega_1 \xrightarrow{f} \Omega_2 @ \mathcal{Q}$ & \mathcal{F}_1 is a σ -field

then $f @ \mathcal{F}_1/\mathcal{Q} \Leftrightarrow f @ \mathcal{F}_1/\mathcal{Q}$.

Proof:

\Rightarrow : trivial

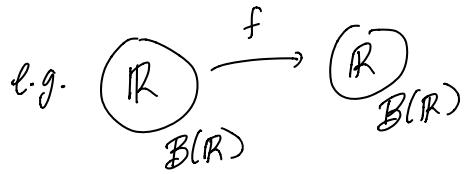
\Leftarrow : Good sets on

$$\mathcal{Y} = \{A \subset \mathbb{R}_1 : f^{-1}(A) \in \mathcal{F}_1\}.$$

$\mathcal{A} \subset \mathcal{Y}$ by assumption & \mathcal{Y} is a σ -field by facts (1), (3), (4).

QED

(3)



f is monotone $\Rightarrow f^{-1}((-\infty, x])$ is an interval $\forall x$

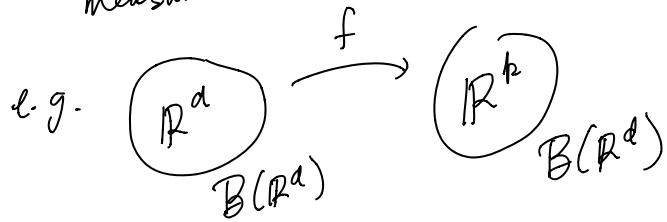
show that $a, b \in f^{-1}((-\infty, x]) \Rightarrow f^{-1}((-\infty, x]) \in \mathcal{B}(\mathbb{R})$, $\forall x$

$\Rightarrow a \leq y \leq b \Rightarrow f(y) \in f^{-1}((-\infty, x])$

$\Rightarrow y \in f^{-1}((-\infty, x])$

could be open or closed $\Leftrightarrow f(\mathcal{B}(\mathbb{R})) / \sigma \{(-\infty, x] : x \in \mathbb{R}\} = \mathcal{B}(\mathbb{R})$

\therefore All monotone funcs are measurable.



f is continuous

$\Leftrightarrow f^{-1}(G)$ is open if open $G \subset \mathbb{R}^k$

$\Rightarrow f(\mathcal{B}(\mathbb{R}^d)) / \text{opens in } \mathbb{R}^k$

$\Leftrightarrow f(\mathcal{B}(\mathbb{R}^d)) / \sigma \{ \text{opens in } \mathbb{R}^k \} = \mathcal{B}(\mathbb{R}^k)$

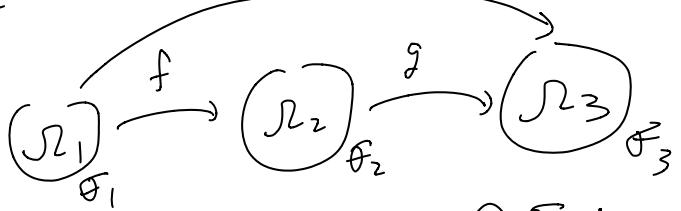
This extends to metric spaces \mathbb{R}_1 & \mathbb{R}_2

so that

$f(\mathcal{B}(\mathbb{R}_1)) / \mathcal{B}(\mathbb{R}_2)$ for all continuous $f: \mathbb{R}_1 \rightarrow \mathbb{R}_2$

Thm (composition of \mathcal{M} is \mathcal{M}). (4)

If



where $f \in \mathcal{F}_1 / \mathcal{F}_2$ & $g \in \mathcal{F}_2 / \mathcal{F}_3$

then $gof \in \mathcal{F}_1 / \mathcal{F}_3$

Proof:

If $B \in \mathcal{F}_3$ then

$$(g \circ f)^{-1}(B) = f^{-1} \circ g^{-1}(B)$$

$$= f^{-1}(g^{-1}(B))$$

$\in \mathcal{F}_2$

QED.

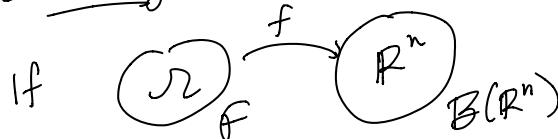
$\in \mathcal{F}_1$

e.g. if $f \in \mathcal{F}/\mathcal{B}(\mathbb{R})$ then

$|f|, f^2, \sin(f) \dots$ are all $\in \mathcal{F}/\mathcal{B}(\mathbb{R})$

by continuity.

Corollary (Just check the coordinates)



works for $\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)$ too

where \mathcal{F} is a σ -field

and $f(w) = (f_1(w), \dots, f_n(w))$

then

$f \in \mathcal{F}/\mathcal{B}(\mathbb{R}^n) \Leftrightarrow$ each f_k is $\in \mathcal{F}/\mathcal{B}(\mathbb{R})$

Proof:

\Rightarrow : follows since the coordinate mappings $\pi_p(x_1, \dots, x_n) = x_p$ are continuous & $f_p = \pi_p \circ f$

\Leftarrow :

$$f^{-1}\left(\underbrace{(a_1, b_1] \times \dots \times (a_n, b_n]}_{\text{bdd rectangle}}\right) = \bigcap_{k=1}^n f_p^{-1}([a_k, b_k]) \in \mathcal{F}$$

$\in \mathcal{F}$ since
 $f_p \in \mathcal{F}/B(\mathbb{R})$

$\in \mathcal{F}$

$\therefore f \in \mathcal{F}/\text{bdd rectangles}$

$\therefore f \in \mathcal{F}/\sigma\langle \text{bdd rectangles} \rangle$

$\underbrace{\quad}_{B(\mathbb{R}^n)}.$

QED.

Thm (Cut & paste over countable & \mathbb{m} pieces)

If $(\mathcal{R}_1, \mathcal{F}_1) \xrightarrow{f} (\mathcal{R}_2, \mathcal{F}_2)$ where

\mathcal{F}_1 & \mathcal{F}_2 are σ -fields and

$$\mathcal{R}_1 = \bigcup_{k=1}^{\infty} A_k \text{ s.t. } f_k, A_k \in \mathcal{F}$$

then

$f \in \mathcal{F}_1/\mathcal{F}_2 \Leftrightarrow f|_{A_k} \in (\mathcal{F}_1 \cap A_k)/\mathcal{F}_2$
for all $k = 1, 2, \dots$

(5)

Proof:

\Rightarrow : suppose $f \in \mathcal{F}_1/\mathcal{F}_2$. Let $B \in \mathcal{F}_2$.

$$w \in f|_{A_k}^{-1}(B) \Leftrightarrow f(w) \in B \text{ & } w \in A_k$$

$$\Leftrightarrow f(w) \in B \text{ & } w \in A_k$$

$$\therefore f|_{A_k}^{-1}(B) = A_k \cap f^{-1}(B) \in \mathcal{F}_1$$

$\in \mathcal{F}_1$ $\in \mathcal{F}_1$

\Leftarrow : suppose $f|_{A_k} \in (\mathcal{F}_1 \cap A_k)/\mathcal{F}_2$. Let $B \in \mathcal{F}_2$.

$$f^{-1}(B) = f^{-1}(B) \cap \mathcal{R}_1$$

$$= f^{-1}(B) \cap \bigcup_k A_k$$

$$= \bigcup_k (f^{-1}(B) \cap A_k)$$

$$= \bigcup_k f|_{A_k}^{-1}(B)$$

$\in \mathcal{F}_1 \cap A_k$ by assumption

$\in \mathcal{F}_1$ since $\mathcal{F}_1 \cap A_k \subset \mathcal{F}_1$
& \mathcal{F}_1 is a σ -field

QED.

Corollary: Piecewise metric

continuous functions are \mathbb{m}
if the "Pieces" are countable &
Borel measurable.

Thm: (just check \mathbb{m} on the range)

If $f: \mathcal{R}_1 \rightarrow \mathcal{R}_2 \subset \mathcal{R}_2$ metric space

and \mathcal{F}_1 is a σ -field on \mathcal{R}_1 , then

$$f \in \mathcal{F}_1/\mathcal{B}(\mathcal{R}_2) \Leftrightarrow f \in \mathcal{F}_1/B(\mathcal{R}_2).$$

(6)

Proof: The borel restriction thm says (7)

$$B(\mathcal{J}_2^\circ) = B(\mathcal{J}_2) \cap \mathcal{J}_2^\circ.$$

∴

$$\begin{aligned} f @ \mathcal{F}_1 / B(\mathcal{J}_2^\circ) &\iff f @ \mathcal{F}_1 / B(\mathcal{J}_2) \cap \mathcal{J}_2^\circ \\ &\iff \underbrace{f^{-1}(B \cap \mathcal{J}_2^\circ)}_{\text{since } f \text{ maps into } \mathcal{J}_2^\circ} \in \mathcal{F}_1, \forall B \in B(\mathcal{J}_2) \\ &\iff f^{-1}(B) \in \mathcal{F}_1, \forall B \in B(\mathcal{J}_2) \\ &\quad \text{since } f \text{ maps into } \mathcal{J}_2^\circ \\ &\iff f @ \mathcal{F}_1 / B(\mathcal{J}_2). \end{aligned}$$

QED.

e.g. $\sin(x) @ B(\mathbb{R}) / B(\mathbb{R})$

$$\iff \sin(x) @ B(\mathbb{R}) / B([-1, 1])$$

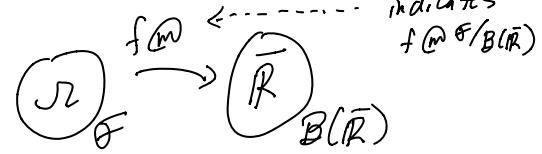
Question:

$$\text{Is } f(x) = \begin{cases} \frac{1}{x}, & x > 0 \\ -\infty, & x = 0 \\ \sin(x), & x < 0 \end{cases} @ B(\mathbb{R}) / B(\overline{\mathbb{R}})$$

Yes since f is metric continuous on countably many measurable pieces.

Notation:

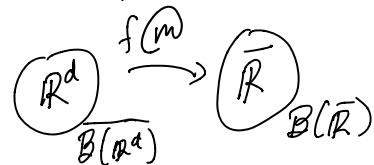
- f is " \mathcal{F} -measurable" if



- f is "Borel measurable" if



- f is "Lebesgue measurable" if



when $\overline{B(\mathbb{R}^d)}$ is the completion w.r.t. Lebesgue measure-

Convention for ∞

- $\infty + x = \infty$ when $x \in (-\infty, \infty]$
- $\infty \cdot 0 = 0$
- $\infty \cdot \infty = \infty$
- $\frac{x}{\infty} = 0$ when $x \in \mathbb{R}$
- $\frac{x}{0}$, $\frac{\pm\infty}{\pm\infty}$, $\infty - \infty$ are not defined.

Thm (closure thm)

(9)

Let (Ω, \mathcal{F}) be a measurable space.

(i) if $\xrightarrow{\substack{f(\omega) \\ g(\omega)}} (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$ then

$f+g, f \cdot g, f/g, \max(f, g), \min(f, g)$,

$f^+, f^-, |f|$ are all $\cap \mathcal{F}/\mathcal{B}(\bar{\mathbb{R}})$

provided they are defined $\forall \omega \in \Omega$

i.e. No $\infty - \infty, \frac{\infty}{\infty} \dots$

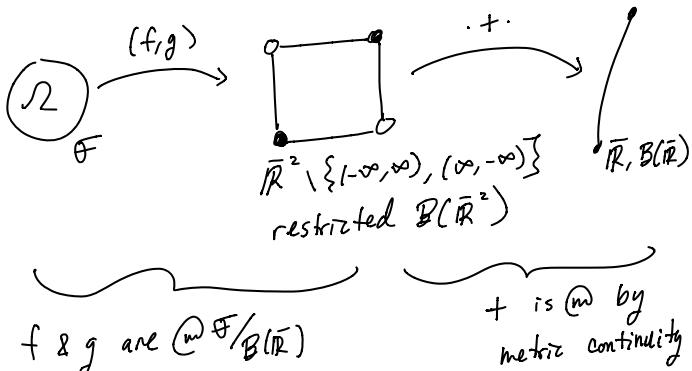
(ii) if $\xrightarrow{\substack{f_i(\omega) \\ f_n(\omega)}} (\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$

then $\sup_n f_n, \inf_n f_n, \limsup_n f_n$

and $\liminf_n f_n$ are $\cap \mathcal{F}/\mathcal{B}(\bar{\mathbb{R}})$.

Proof:

(i) Just show $f+g \dots$ the others are similar.
since $f(\omega) + g(\omega)$ is defined $\forall \omega \in \Omega$ we have



$\Leftrightarrow (f, g)$ is $\cap \mathcal{F}/\mathcal{B}(\bar{\mathbb{R}}^2)$
by "just check cords"

$\Leftrightarrow (f, g)$ is \cap on the
restricted space, by
"just check the range"

$\therefore f+g \in \mathcal{F}/\mathcal{B}(\bar{\mathbb{R}})$ by "composition of
 \cap is \cap ".

To show (ii) just notice

(10)

$$(\sup_n f_n)^{-1}([-\infty, c]) = \left\{ \omega : \sup_n f_n(\omega) \leq c \right\}$$

$$= \bigcap_{n=1}^{\infty} \left\{ \omega : f_n(\omega) \leq c \right\}$$

Note: this will not be true for ∞

$\in \mathcal{F}$ since \mathcal{F} is a σ -field.

$$\therefore \sup_n f_n \in \mathcal{F}/\{[-\infty, c] : c \in \bar{\mathbb{R}}\}$$

$$\therefore \sup_n f_n \in \mathcal{F}/\underbrace{\{[-\infty, c] : c \in \bar{\mathbb{R}}\}}_{= \mathcal{B}(\bar{\mathbb{R}})}.$$

For the others

$$\inf_n f_n = -\sup_n (-f_n)$$

$$\limsup_{n \rightarrow \infty} f_n = \lim_{m \rightarrow \infty} \sup_{n \geq m} f_n = \inf_m \sup_{n \geq m} f_n$$

decreases as $m \rightarrow \infty$

$$\liminf_{n \rightarrow \infty} f_n \dots \text{similar.}$$

Q.E.D.

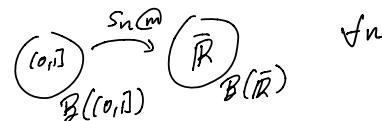
E.g. Coin flip model from lecture 1.

$$S_n(\omega) = \sum_{k=1}^n X_k(\omega) \text{ maps } (0,1] \rightarrow \bar{\mathbb{R}}$$

Since S_n is constant over intervals $(\frac{i-1}{2^n}, \frac{i}{2^n}]$
we have $S_n^{-1}([-\infty, x]) = \text{finite disjoint union of dyadic}$

union of dyadiques

By "generators are enough"



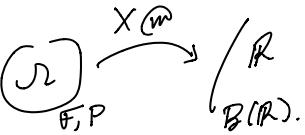
$\therefore \limsup_n \frac{S_n}{\sqrt{n \log n}}$ is $\cap \mathcal{B}((0,1])/\mathcal{B}(\bar{\mathbb{R}})$

$\therefore \left\{ \limsup_n \frac{S_n}{\sqrt{n \log n}} = 1 \right\} \in \mathcal{B}((0,1])$ since
it is the pre-image of $\{1\} \in \mathcal{B}(\bar{\mathbb{R}})$.

Random variables, induced measures and C-d.f.s

II

Def: X is a random variable if there exists a probability space (Ω, \mathcal{F}, P) where $X: \Omega \rightarrow \mathbb{R}$ s.t. $X \in \mathcal{F}/B(\mathbb{R})$.
 X is an extended random variable if $X: \Omega \rightarrow \bar{\mathbb{R}}$ & $X \in \mathcal{F}/B(\bar{\mathbb{R}})$.

Picture: 

we write $X(\omega)$ instead of $f(\omega)$ to indicate (Ω, \mathcal{F}) has a probability measure attached.
 Think of P as modeling a random draw $\omega \in \Omega$ & $X(\omega)$ as a "variable" or "label" associated with each $\omega \in \Omega$.

Since $X \in \mathcal{F}/B(\mathbb{R})$ it makes sense to talk about quantities like:

$$P(X=1) = P(X^{-1}(\{1\}))$$

$\underbrace{\quad}_{\in B(\mathbb{R})}$

$$P(X \leq x) = P(X^{-1}((-\infty, x]))$$

$\underbrace{\quad}_{\in \mathcal{F}}$

$$P(X \in \mathbb{Q}) = P(X^{-1}(\mathbb{Q})).$$

(12)

Def: If X is a random variable defined on (Ω, \mathcal{F}, P) , the distribution of X (also called the induced probability measure) is a set function $PX^{-1}: B(\mathbb{R}) \rightarrow [0, 1]$ given by

$$PX^{-1}(B) := P(X^{-1}(B)) = P(X \in B).$$

More generally if



where $\mathcal{F}_1, \mathcal{F}_2$ are σ -fields & $(\Omega_1, \mathcal{F}_1, \mu)$ is a measure then

$$\mu f^{-1}(F) := \mu(f^{-1}(F)), \quad \forall F \in \mathcal{F}_2$$

\nwarrow induced measure on $(\Omega_2, \mathcal{F}_2)$.

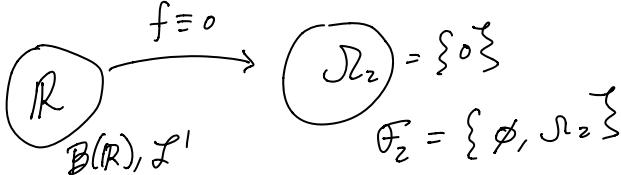
Thm: In the setup above μf^{-1} is a measure on $(\Omega_2, \mathcal{F}_2)$

Proof:
 Follows immediately since pre-images & set operations commute. QED

Since $\mu f^{-1}(\Omega_2) = \mu(\Omega_1)$ we have the following facts:

- $\mu(\Omega_1) = 1 \Rightarrow \mu f^{-1}(\Omega_2) = 1$
i.e. PX^{-1} is a probability measure
- $\mu(\Omega_1) < \infty \Rightarrow \mu f^{-1}(\Omega_2) < \infty$
- Warning:
 μ is a σ -finite measure
 $\nrightarrow \mu f^{-1}$ is a σ -finite measure

e.g.



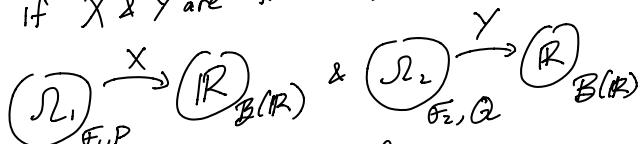
f is $B(R)/F_2$ measurable and L^1 is σ -finite but $L^1 f^{-1}$ is not σ -finite since

$$L^1 f^{-1}(\phi) = 0 \text{ and}$$

$$L^1 f^{-1}(\Omega_2) = f^{-1}(R) = \infty.$$

Notice that even if two r.v.s X & Y are defined on different probability spaces the induced distributions are both on $(R, B(R))$.

Def: If X & Y are two r.v.s s.t.



we write $X \sim Y$ or $X = Y$ if

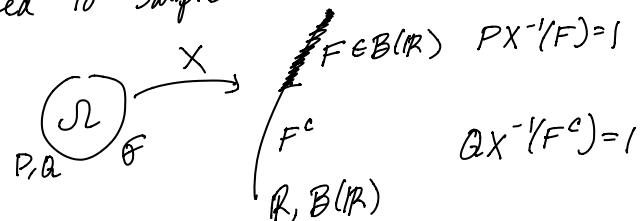
$$PX^{-1} = QY^{-1} \text{ over } B(R).$$

(13)

In the simple statistics setup (14)
there are two possible measures P & Q , both defined on (Ω, \mathcal{F}) .
 $w \in \Omega$ is picked at random either from P or Q & the "observer" only gets to see the value of $X(w)$ (i.e. the data).

The observer then tries to figure out which P or Q w was sampled from.

Note: using the notation from HWK3 if $PX^{-1} \perp QX^{-1}$ then the observer will know exactly which P or Q was used to sample w



- if $X(w) \in F$ then w was drawn from P .
- if $X(w) \in F^c$ then w was drawn from Q .

Def: The distribution function (sometimes called cumulative distribution function.. CDF) for a random variable X defined on (Ω, \mathcal{F}, P) is the function $F: \mathbb{R} \rightarrow [0, 1]$ defined as

$$\begin{aligned} F(t) &= P(X \leq t), \quad \forall t \in \mathbb{R} \\ &= PX^{-1}(-\infty, t] \end{aligned}$$

Thm: If X & Y are two r.v.s with c.d.f.s F_X & F_Y , respectively then

$$X \stackrel{d}{=} Y \Leftrightarrow F_X(t) = F_Y(t) \quad \forall t \in \mathbb{R}.$$

Proof:

\Rightarrow : trivial when taking $B = (-\infty, t]$

$$\text{since } X^{-1}(-\infty, t] = \{X \leq t\}$$

\Leftarrow :

$$F_X(t) = F_Y(t) \quad \forall t \in \mathbb{R}$$

$$\Rightarrow P(X^{-1}(-\infty, t]) = P(Y^{-1}(-\infty, t]) \quad \forall t \in \mathbb{R}.$$

$$\Rightarrow P(X^{-1}) = P(Y^{-1}) \text{ on } \Omega = \{(-\infty, t] : t \in \mathbb{R}\}$$

$$\xrightarrow{\text{II-ning}} \Rightarrow P(X^{-1}) = P(Y^{-1}) \text{ on } \sigma(\Omega) = \mathcal{B}(\mathbb{R}).$$

$\nwarrow_{\text{II-sys}}$

QED.

Thm (Properties of c.d.f.s)

Let $F(t) = P(X \leq t)$ for a r.v. X defined on (Ω, \mathcal{F}, P) . Then

$$(I) \quad F(x) \leq F(y), \quad \forall x \leq y$$

$$(II) \quad \lim_{x \downarrow y} F(x) = F(y)$$

$$(III) \quad \lim_{x \rightarrow \infty} F(x) = 1 \quad \& \quad \lim_{x \rightarrow -\infty} F(x) = 0$$

(15)

Proof:

(I): clear since $A \subset B \Rightarrow P(X^{-1}(A)) \leq P(X^{-1}(B))$

(II): $\{X \leq x\} \downarrow \{X \leq y\} \text{ as } x \downarrow y$

$$\therefore P(X \leq x) \downarrow P(X \leq y).$$

Warning: it is not true that
 $\{X < x\} \downarrow \{X < y\}$
e.g. take $x = \frac{1}{n}$, $y = 0$, $X \geq 0$ & $P(X=0) > 0$

(III): Follows since

$$\{X \leq x\} \uparrow \mathbb{R} \quad \& \quad \{X \leq x\} \downarrow \emptyset.$$

QED.

It turns out properties (I), (II), (III)
are characterizing features of
c.d.f.s i.e. if $F: \mathbb{R} \rightarrow [0, 1]$ satisfies

(I), (II), (III) then \exists a r.v. X
s.t. $F(t) = P(X \leq t)$.

In fact, this will be the main tool
we use to show the existence of
an infinite sequence of indep. r.v.s
all with a specified distribution.

Dfn: if $F: \mathbb{R} \rightarrow [0, 1]$ satisfies
(I), (II) & (III) of the above thm
define

$$F^{-1}(u) := \inf \{x \in \mathbb{R} : F(x) \geq u\}$$

(16)

Fact 1: The inf in F^{-1} is attained. (17)

(18)