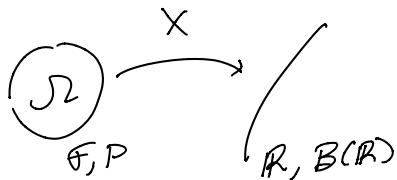


# Lecture 8: Measurable functions, Random variables and distribution functions

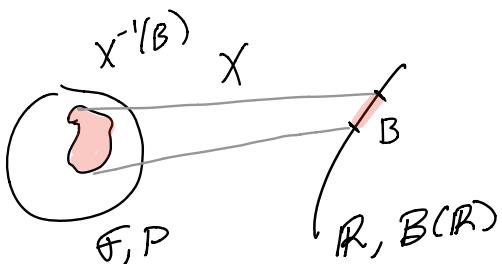
In this lecture we will start by developing the measure theoretic notion of measurable functions & define random variables  $X$  as measurable functions  $X: \Omega \rightarrow \mathbb{R}$  where  $(\Omega, \mathcal{F}, P)$  is a prob. space.



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  $\sigma$ -fields ... just collections of sets.

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

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

(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  $\sigma$ -field

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

Proof:

$\Rightarrow$ : trivial

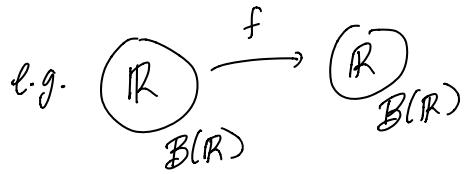
$\Leftarrow$ : Good sets on

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

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

QED

(3)



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

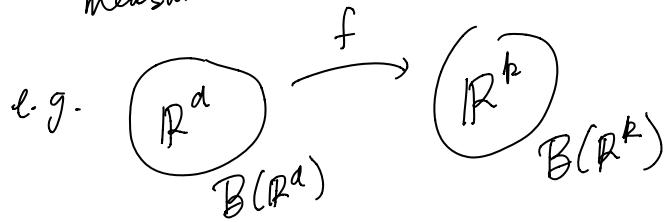
show that  $a, b \in f^{-1}((-\infty, x]) \Rightarrow f^{-1}((-\infty, x]) \in \mathcal{B}(\mathbb{R})$ ,  $\forall x$

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

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

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  $G \subset \mathbb{R}^k$

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

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

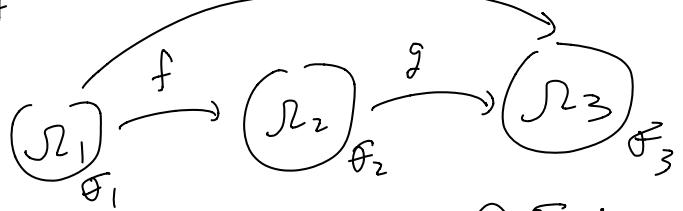
This extends to metric spaces  $\mathbb{R}_1$  &  $\mathbb{R}_2$

so that

$f(\mathcal{B}(\mathbb{R}_1)) / \mathcal{B}(\mathbb{R}_2)$  for all metric 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.

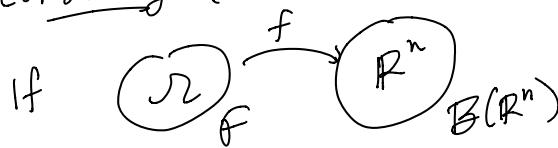
$\in \mathcal{F}_1$

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  $\sigma$ -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:

(5)

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

$\Leftarrow$ :

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

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

$\therefore f \in \mathcal{F}/\underbrace{\sigma(\text{rectangles})}_{B(\mathbb{R}^n)}$ .

QED.

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

If  $\mathcal{R}_1 \xrightarrow[f]{\mathcal{F}_1} \mathcal{R}_2$  where

$\mathcal{F}_1, \mathcal{F}_2$  are  $\sigma$ -fields and

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

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  $k = 1, 2, \dots$

Proof:

(6)

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

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

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

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

$\Leftarrow$ : 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  $\sigma$ -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  $\sigma$ -field on  $\mathcal{R}_1$ , then

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

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

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

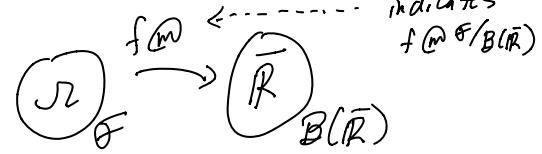
Question:

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

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

Notation:

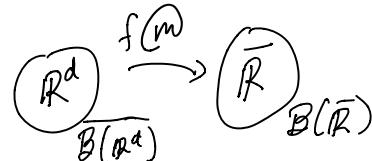
- $f$  is " $\mathcal{F}$ -measurable" if



- $f$  is "Borel measurable" if



- $f$  is "Lebesgue measurable" if

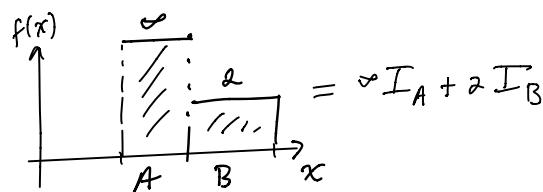


when  $\overline{B(\mathbb{R}^d)}$  is the completion w.r.t. Lebesgue measure-

Convention for  $\infty$

- $\infty + x = \infty$  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.

Note: we use the convention  $\infty \cdot 0 = 0$  so



# Thm (closure thm)

(9)

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

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

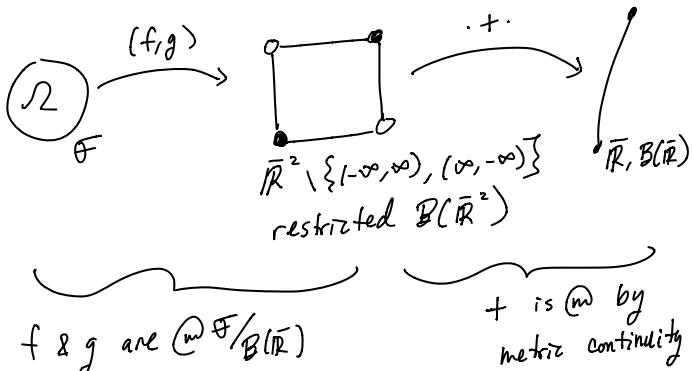
$f+g, f \cdot g, f/g, \max(f, g), \min(f, g), f^+, f^-$ ,  $|f|$  are all  $\cap \mathcal{F}/\mathcal{B}(\bar{\mathbb{R}})$   
provided they are defined  $\forall \omega \in \Omega$   
i.e. No  $\infty - \infty, \frac{\infty}{\infty} \dots$

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

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

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

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

Note: this will not be true for  $\infty$  since  $f_n \in \mathcal{F}/\mathcal{B}(\bar{\mathbb{R}})$ .

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

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

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

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  $m \rightarrow \infty$

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

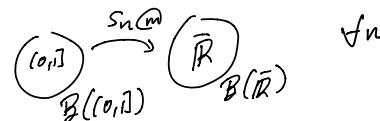
Q.E.D.

E.g. Coin flip model from lecture 1.

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

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

"generators are enough"



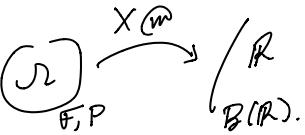
$\therefore \limsup_n \frac{S_n}{\sqrt{n \log n}}$  is  $\cap \mathcal{B}((0, 1))/\mathcal{B}(\bar{\mathbb{R}})$

$\therefore \left\{ \limsup_n \frac{S_n}{\sqrt{n \log n}} = 1 \right\} \in \mathcal{B}((0, 1))$  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  $(\Omega, \mathcal{F}, P)$  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  $(\Omega, \mathcal{F})$  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  $(\Omega, \mathcal{F}, P)$ , 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  $\sigma$ -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  $(\Omega_2, \mathcal{F}_2)$ .

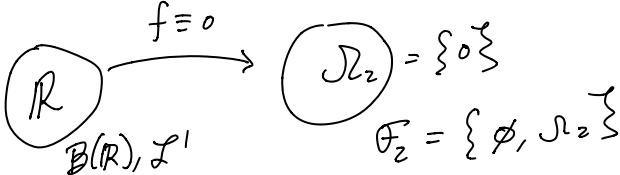
Thm: In the setup above  $\mu f^{-1}$  is a measure on  $(\Omega_2, \mathcal{F}_2)$

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

Since  $\mu f^{-1}(\mathcal{R}_2) = \mu(\mathcal{R}_1)$  we have the following facts:

- $\mu(\mathcal{R}_1) = 1 \Rightarrow \mu f^{-1}(\mathcal{R}_2) = 1$   
i.e.  $PX^{-1}$  is a probability measure
- $\mu(\mathcal{R}_1) < \infty \Rightarrow \mu f^{-1}(\mathcal{R}_2) < \infty$
- Warning:  
 $\mu$  is a  $\sigma$ -finite measure  
 $\nrightarrow \mu f^{-1}$  is a  $\sigma$ -finite measure

e.g.



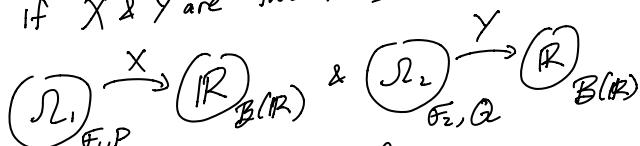
$f$  is  $B(R)/F^1$  measurable and  $\mathbb{Z}'$  is  $\sigma$ -finite but  $\mathbb{Z}'f^{-1}$  is not  $\sigma$ -finite since

$$\mathbb{Z}'f^{-1}(\phi) = 0 \text{ and}$$

$$\mathbb{Z}'f^{-1}(\mathcal{R}_2) = \mathbb{Z}'(R) = \infty.$$

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  $X \stackrel{d}{=} Y$  if

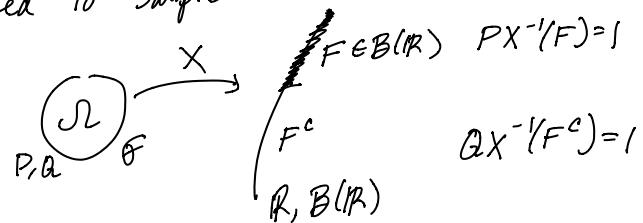
$$PX^{-1} = QY^{-1} \text{ over } B(R).$$

(13)

In the simple statistics setup (14)  
there are two possible measures  $P$  &  $Q$ , both defined on  $(\mathcal{R}, \mathcal{F})$ .  
 $w \in \mathcal{R}$  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.. c.d.f) for a random variable  $X$  defined on  $(\mathcal{R}, \mathcal{F}, P)$  is the function  $F: \mathbb{R} \rightarrow [0, 1]$  defined as

$$F(t) = P(X \leq t), \quad \forall t \in \mathbb{R}.$$

$$= PX^{-1}(-\infty, t]$$

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  $B = (-\infty, t]$

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

$\Leftarrow$ :

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

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

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

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

AED.

(15)

Proof:

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

(II):

$$\{X \leq x\} \downarrow \{X \leq y\} \text{ as } x \downarrow y$$

which follows since if  $x_n \downarrow y$  as  $n \rightarrow \infty$   
 then clearly  $\{X \leq y\} \subset \bigcap_{n=1}^{\infty} \{X \leq x_n\}$  &  
 $\omega \in \bigcap_{n=1}^{\infty} \{X \leq x_n\} \Rightarrow X(\omega) \leq x_n \xrightarrow{n \rightarrow \infty} X(\omega) \leq \lim_n x_n = y$

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

**Warning:** It is not true that  
 $\{X < x\} \downarrow \{X < y\} \text{ as } x \downarrow y$   
 since  $X(n) < x_n \not\Rightarrow X(n) < \lim x_n = y$   
 e.g. take  $X(n) = 0, x_n = \frac{1}{n}, y = 0$

(III): Follows since

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

AED.

It turns out properties (I), (II), (III)

are characterizing properties 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.

(16)

Deg.: If  $F: \mathbb{R} \rightarrow [0,1]$  satisfies  
(I), (II) & (III) of the above then  
define

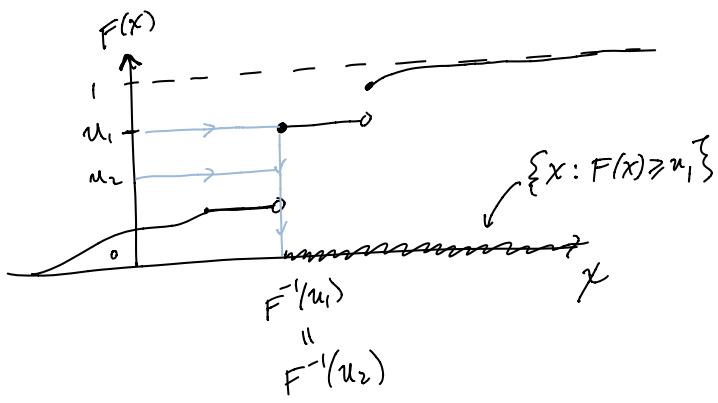
$$F^{-1}(u) := \inf \{x \in \mathbb{R} : F(x) \geq u\}$$

↗ often called the  
inverse c.d.f. or the  
quantile function.

for  $u \in (0,1)$ .

Note:  $F^{-1}: (0,1) \rightarrow \mathbb{R}$  is well defined  
since  $\{x : F(x) \geq u\} \neq \emptyset$  when  $0 < u < 1$   
by (III) & (I).

Picture:



There are two important facts about  $F$  &  $F^{-1}$  which are useful to remember.

Suppose  $U$  is a r.v. which is uniform on  $(0,1)$ , i.e.  $\exists (\Omega, \mathcal{F}, P)$  &  $\xrightarrow[\mathcal{F}, P]{u \sim} (0,1)$

$$\text{s.t. } P(U \in B) = \mathcal{J}'(B).$$

(17)

For example, digit coin flip model: (18)  
 $\Omega = \{0,1\}$ ,  $\mathcal{F} = \mathcal{B}(\{0,1\})$ ,  $P = \mathcal{J}'|_{\{0,1\}}$

where

$$u(w) = \begin{cases} w & \text{if } w < 1 \\ 1 & \text{if } w = 1 \end{cases}$$

is (17) by cut & paste &

$$P(U \in B) = P(B) = \mathcal{J}'(B)$$

↗ since  $P(U = 1) = 0$

Now the two important facts are --  
if  $F$  is the c.d.f. of a r.v.  $X$  then

$$(a) \quad F^{-1}(u) \stackrel{d}{=} X$$

$$(b) \quad F(X) = U \text{ when } F \text{ is continuous.}$$

Lemma (switching formula):

If  $F: \mathbb{R} \rightarrow [0,1]$  satisfies (I), (II) & (III)  
above then

$$(*) \quad F(x) \geq u \iff x \geq F^{-1}(u)$$

$\forall u \in (0,1) \text{ & } x \in \mathbb{R}$ .

Proof:

A useful restatement of (\*) is simply that

$$(**) \quad \{x : F(x) \geq u\} = [F^{-1}(u), \infty).$$

To show (\*\*), notice that

$\{x : F(x) \geq u\}$  must be an interval of the form  $[F^{-1}(u), \infty)$  or  $[F^{-1}(u), \infty)$  since

$$F^{-1}(u) \stackrel{\text{def}}{=} \inf \{x : F(x) \geq u\}$$

and

$$x' \in \{x : F(x) \geq u\} \xrightarrow{\text{by (I)}} \begin{cases} x'' \in \{x : F(x) \geq u\} \\ \forall x'' > x' \end{cases}$$

∴ to show  $\{x: F(x) \geq u\} = [F^{-1}(u), \infty)$  just (1a)  
prove the inf of the LHS is attained.

Let  $x_p \in \{x: F(x) \geq u\}$  s.t.  $x_p \downarrow \underbrace{F^{-1}(u)}_{=\inf \text{ of LHS}}$

∴  $F(\lim_p x_p) \stackrel{(II)}{=} \lim_p F(x_p) \geq u$   
↑ since  $x_p$  is  
in  $\{x: F(x) \geq u\}$

∴  $F^{-1}(u) = \lim_k x_k \in \{x: F(x) \geq u\}$   
so the inf is attained.

(Note we used III implicitly to show  
 $F^{-1}(u)$  is well defined and a r.v.)

QED

Lemma (c.d.f sandwich)

If  $F: \mathbb{R} \rightarrow [0,1]$  satisfies (I), (II) & (III)

then  $\forall u \in (0,1)$

$$F(F^{-1}(u)-) \leq u \leq F(F^{-1}(u)).$$

Proof:

$u \leq F(F^{-1}(u))$  holds since  $F^{-1}(u) \in \{x: F(x) \geq u\}$ .

The contrapositive of the switching formula is

$$F(x) < u \Leftrightarrow x < F^{-1}(u)$$

$$\therefore F(F^{-1}(u)-) = \lim_{\substack{x \uparrow F^{-1}(u)}} \underbrace{F(x)}_{< u} = u$$

these satisfy  $x < F^{-1}(u)$   
so  $F(x) < u$  by switch

QED

Maybe the best way to remember these

$$\{x: F(x) \geq u\} = [F^{-1}(u), \infty)$$

$$F(F^{-1}(u)-) - u \leq \underbrace{F(F^{-1}(u)+)}_{= F(F^{-1}(u)) \text{ by (II)}} \leq u$$

Thm (c.d.f representation)

if  $F: \mathbb{R} \rightarrow [0,1]$  satisfies (I), (II) & (III)  
then  $\exists$  a r.v.  $X$  on a prob space  $(\Omega, \mathcal{F}, P)$

s.t.

$$(1) \quad P(X \leq x) = F(x) \quad \forall x \in \mathbb{R}$$

and, moreover, any r.v.  $U$  which is  
uniformly distributed on  $(0,1)$  satisfies

$$(2) \quad F^{-1}(U) \stackrel{D}{=} X$$

$$(3) \quad P(F(X) \leq u) = u \quad \text{for all } u \in (0,1).$$

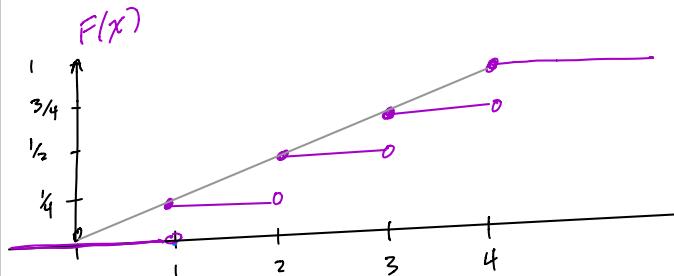
If, in addition,  $F$  is continuous then

$$(4) \quad F(X) = U.$$

Here is the picture of why we don't  
always get (4).

Suppose  $P(X=i) = \frac{1}{4}$  for  $i=1, 2, 3, 4$   
So  $X$  is uniformly distributed on  $\{1, 2, 3, 4\}$

The c.d.f of  $X$  is



∴  $F(X)$  assigns  $\frac{1}{4}$  prob  
to  $\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ .

$$\text{i.e. } P(F(X) \leq 0.25) = 0.25$$

$$\text{but } P(F(X) \leq 0.3) = 0.25 \leq 0.3$$

(21)

Proof:

First note that  $F^{-1}$  is C<sup>m</sup> by monotonicity.Let  $U$  be a r.v. uniform on  $(0,1)$ .

$$\begin{aligned} \therefore P(F^{-1}(U) \leq x) &= P(U \leq F(x)) \\ &\quad \text{by switching lemma} \\ &= F'(F(x)) \quad U \text{ is uniform} \\ &= F(x) \end{aligned}$$

Now set  $X := F^{-1}(U)$  to get (1) & (2) using  
1-1 uniqueness.

For (3),

$$\begin{aligned} P(F(X) \leq u) &= P(F(F^{-1}(U)) \leq u) \\ &\quad \text{since } X = F^{-1}(U) \\ &\leq P(U \leq u) \\ &\quad \text{since } U \leq F(F^{-1}(u)) \leq u \\ &\quad \uparrow \\ &\quad \text{by c.d.f sandwich} \\ &= u, \quad \forall u \in (0,1) \end{aligned}$$

If  $F$  is continuous c.d.f sandwichgives  $F(F^{-1}(u)-) = U = F(F^{-1}(u))$  so

$$P(F(X) \leq u) = P(U \leq u)$$

which shows (4)

QED