Background

$\mathbb{N} = \{1, 2, 3, \cdots\}$ — natural #'s

$Z = \{0, \pm 1, \pm 2, \cdots\}$ — integers

sets which can be put into a 1-1 correspondence with $\mathbb{N}$ are <u>countably infinite</u>. sets which have a finite # of elements are <u>countably finite</u>. These two types are the <u>countable</u> sets. Usually countable will mean countably infinite.

Z is countable. So is Z^2 & so is Q = set of <u>rationals</u>

$$\frac{m}{n} \quad m \leftarrow \text{integers}, \quad n \neq 0$$

If a set is not countable it is <u>uncountable</u>. An example would be $[0,1]$ or $\mathbb{R}$.

Let $a_1, a_2, \cdots$ be a sequence. Consider

$$\sup\{a_n : n \geq N\} = \text{lub}\,\{a_n : n \geq N\}$$

(this is either finite or $+\infty$).

Now define

$$\limsup_{n\to\infty} a_n = \lim_{N\to\infty} \sup\{a_n : n \geq N\}$$

This is either finite or $\pm\infty$. We will denote it by $\overline{\lim}\, a_n$. In the same way

$$\underline{\lim}\, a_n = \liminf_{n\to\infty} a_n = \lim_{N\to\infty} \underset{\text{glb}}{\inf}\{a_n : n \geq N\}$$

This is either finite or $\pm\infty$.

$$\lim_{n\to\infty} a_n = a \Leftrightarrow \underline{\lim}\, a_n = \overline{\lim}\, a_n = a$$

Also it is always true $\underline{\lim}\, a_n \leq \overline{\lim}\, a_n$.

This can be extended to $a(t)$, $-\infty < t < \infty$ in the obvious way so we have

$$\limsup_{t\to\infty} a(t)$$

$\underset{\sim}{a}_n \to \underset{\sim}{a}$ if $\forall \epsilon > 0 \quad \exists N$ such that $n \geq N \Rightarrow |\underset{\sim}{a}_n - \underset{\sim}{a}| \leq \epsilon$

$a_n \to a \Rightarrow a_{n_k} \to a$ as $k \to \infty$ where $\{a_{n_k}\}$ is a subsequence

By convention $n_1 < n_2 < \cdots$

On the other hand, if every $a_{n_k} \to a \Rightarrow a_n \to a$

<u>Prop</u> Let $\{a_n\}$ be a sequence of #'s.
If every subsequence $\{a_{n_k}\}$ has a further subsequence

$$a_{n_{k_j}} \to a$$

then $a_n \to a$.

$a_n \to a$ if for every <u>rational</u> $\epsilon_r > 0$ $\exists$ N st $n \geq N$
$\Rightarrow |a_n - a| \leq \epsilon_r$

$a_n \uparrow a$ if $a_1 \leq a_2 \leq \cdots$ and $a_n \to a$
$a_n \downarrow a$ if $a_1 \geq a_2 \geq \cdots$ "

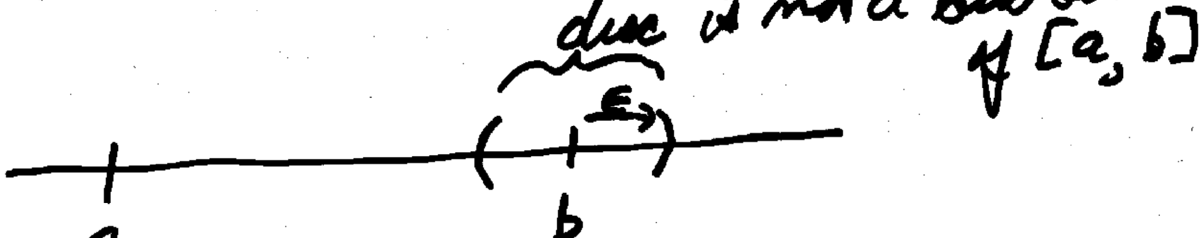
<u>Fact</u> $\lim_{x \to a} g(x) = l$ if $\forall x_n \to a$
$g(x_n) \to l$. In fact this is
$\Leftrightarrow$ if $\forall x_n \uparrow a$ or $x_n \downarrow a$ we have
$g(x_n) \to l$.

Another way of approaching $a_n \to a$.
$a_n \to a$ if for every $\epsilon_r > 0$ only
a finite number of the statements
$|a_n - a| > \epsilon_r$ hold.

(a, b) is an open interval

$\{\underset{\sim}{x} \mid |\underset{\sim}{x} - \underset{\sim}{a}| < \epsilon\}$ is an open disc (the inside of a circle/sphere of radius ϵ). A set $B \subset \mathbb{R}^k$ is open if for every $\underset{\sim}{a} \in B$ there is a disc $\{\underset{\sim}{x} \mid |\underset{\sim}{x} - \underset{\sim}{a}| < \epsilon\} \subset B$.

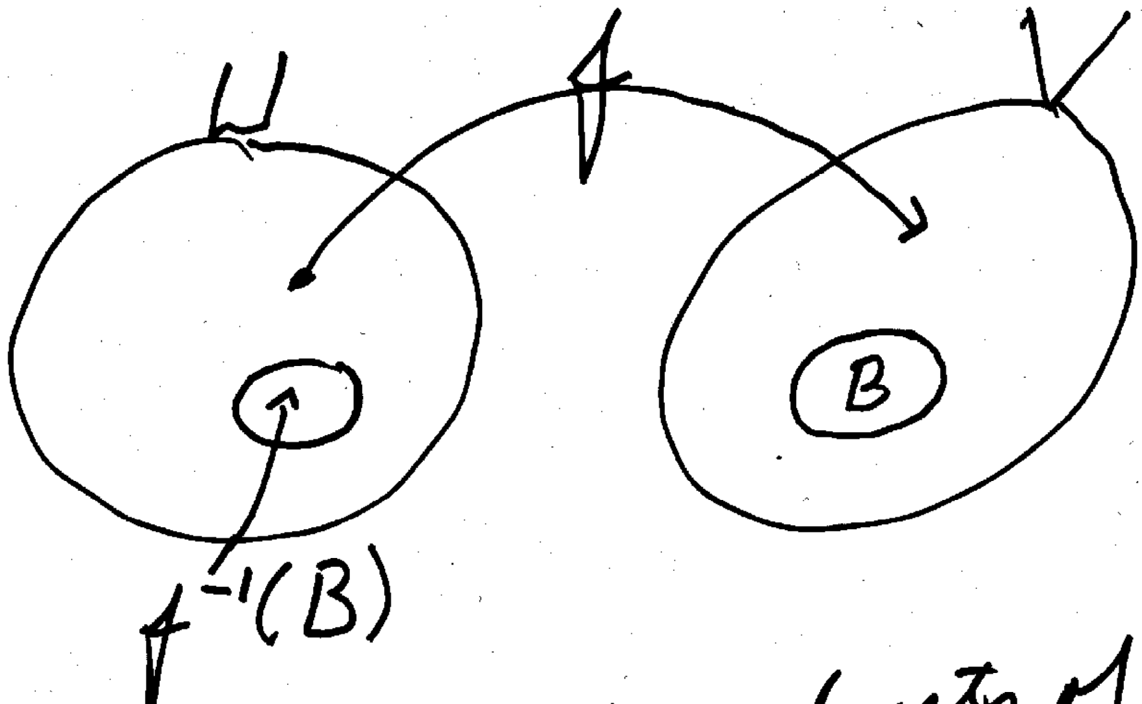
$[a, b]$ is clearly not open



Suppose $B \subset \mathbb{R}$ is open. Then B is a countable union of <u>disjoint</u> open intervals.

The rationals are <u>dense</u> in $\mathbb{R}^{(k)}$

$f: U \to V$



Let $\mathcal{G}$ = set of subsets of V.

Then $f^{-1}(\mathcal{G}) = \{ f^{-1}(B) \mid B \in \mathcal{G} \}$

Let $\Omega = \{\omega\}$ be a finite set. The collection of all subsets of Ω is called the power set on Ω. This power set includes Ω & $\emptyset$. The # of elements in this power set is $2^{|\Omega|}$.

Now suppose $\Omega = \mathbb{N}$. Would the power set be countable? No! $\therefore$ the # of sequences of the type $1, 1, 0, 1, 0, \cdots$ is uncountable.

<u>Def'n</u> $\mathcal{F}$ is a σ-field of subsets of Ω if

(i) $\phi \in \mathcal{F}$ (ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

(iii) $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$

<u>Recall</u>
$$\left.\begin{aligned} \Big(\bigcup_t A_t\Big)^c &= \bigcap_t A_t^c \\ \Big(\bigcap_t A_t\Big)^c &= \bigcup_t A_t^c \end{aligned}\right\} \text{de Morgan}$$

<u>Fact</u> If $\{\underset{\sim}{F}_t, t \in T\}$ is a collection of σ-fields then $\bigcap_{t \in T} \underset{\sim}{F}_t$ is a σ-field.

index set

Let $\mathcal{C}$ be a collection of subsets of Ω. Clearly

$$\mathcal{C} \subset \underbrace{\text{power set}}_{\sigma\text{-field}}$$

$\bigcap_{\text{all } \underset{\sim}{F} \supset \mathcal{C}} \underset{\sim}{F}$ is the σ-field generated by $\mathcal{C}$ & is denoted by $\sigma(\mathcal{C})$. It is the smallest σ-field which includes $\mathcal{C}$.

On $\mathbb{R}^k$ the σ-field generated by the open sets is called the Borel σ-field & denoted by $\mathcal{B}_k$. This σ-field is very rich & includes open/closed sets etc...

Now consider

$$X : \Omega \to \mathbb{R}$$

Let $\mathcal{F}$ be a σ-field on Ω.
If $X^{-1}(\mathcal{B}) \subset \mathcal{F}$ then we say that X is measurable wrt $\mathcal{F}$ & $\mathcal{B}$.

$(\Omega, \mathcal{F}, P)$ is called a probability space.

$\uparrow$ sample space

$X_n \xrightarrow{wp1} X$ or $X_n \xrightarrow{as} X$

"with prob 1" "almost surely"

if $P(\underbrace{\lim_{n\to\infty} X_n}_{rv} = X) = 1$. This is $\Leftrightarrow$

$$P(\{\omega \mid X_n(\omega) \not\to X(\omega)\}) = 0$$

Suppose for each rational $\epsilon_r > 0$

$$P(\underbrace{|X_n - X| > \epsilon_r \quad io}_{A_{\epsilon_r}}) = 0$$

A_{ϵ_r} {The event that an ∞ # of $|X_n - X| > \epsilon_r$. io stands for "infinitely often".

Since $P\left(\bigcup_{\epsilon_r > 0} A_{\epsilon_r}\right) \le \sum_{\epsilon_r > 0} P(A_{\epsilon_r}) = 0$ we

conclude that there is an event A^c with $P(A^c)=1$

such that for each $\epsilon_r > 0$

$$P(\{\omega \in A^c \mid \text{a finite \# of } |X_n(\omega) - X(\omega)| > \epsilon_r\}) = 1$$

Note A^c does not depend on ϵ_r. It then follows

$$X_n \xrightarrow{as} X$$

More on convergence

Events $A_1, A_2, \cdots$ $(\infty \#)$

Borel Cantelli Lemma

(i) $P(A_n \text{ io}) = 0$ if $\sum P(A_k) < \infty$

(ii) A's ind & $\sum PA_k = \infty$ then $P(A_n \text{ io}) = 1$

$$X_n \xrightarrow{as} X \quad (X_n \xrightarrow{wp1} X)$$

Suppose
$$\sum P(|X_n - X| > \epsilon) < \infty$$

$$\overset{BCL}{\Rightarrow} P(|X_n - X| > \epsilon_r \text{ io}) = 0$$

$\Rightarrow$ All X_n are within ϵ_r of X wp1

Now let B_r be this event st $P(B_r) = 1$. Set $B = \bigcap_r B_r \Rightarrow P(B) = 1$ & on B all X_n are within ϵ of X for any $\epsilon > 0$.

$$\Rightarrow X_n \xrightarrow{as} X$$

Lemma Let B_r be st $P(B_r) = 1$, $\forall r \in Q$ Then $P(\bigcap_r B_r) = 1$

Proof $P((\bigcap_r B_r)^c)$

$$= P\left(\bigcup_r B_r^c\right) \leq \sum_r P(B_r^c) = 0$$

$$\Rightarrow P((\bigcap_r B_r)^c) = 0 \quad \Rightarrow P(\bigcap_r B_r) = 1$$

qed

$$X_n \xrightarrow{P} X \quad \left(P(|X_n - X| > \epsilon) \to 0 \right)$$

$$\Rightarrow \exists\ X_{n_k} \xrightarrow{as} X$$

example Suppose $X_n \xrightarrow{P} X$ & $|X_n| \overset{as}{\leq} W$ where $E(W) < \infty$. Then $E(X_n) \to E(X)$

Proof Set $a_n = E(X_n)$ & $a = E(X)$. Let a_{n_k} be any subsequence of $\{a_n\}$. Now look at $X_{n_k} \xrightarrow{P} X$. Then there is a subsequence of this subsequence, say $X_{n_{k_i}}$ such that $X_{n_{k_i}} \xrightarrow{as} X$. Now use the

DCT to get

$$E(X_{m_{k_i}}) \to E(X)$$

That is

$$a_{m_{k_i}} \to a \quad .$$

$\therefore \quad a_m \to a$

$\underline{\underline{qed}}$

A bit more on σ-fields

$(a, b) \leftarrow$ open interval $\subset \mathbb{R}$

<u>open</u> subset of $\mathbb{R}$

closed set = complement of an open set

open set on $\mathbb{R}$ = <u>countable</u> union of <u>disjoint</u> open intervals

open set on $\underbrace{\mathbb{R}^n}_{n>1}$ = countable union of open $\underbrace{\text{"intervals"}}_{\{\underset{\sim}{x} \mid \underset{\sim}{a} < \underset{\sim}{x} < \underset{\sim}{b}\}}$

On $\mathbb{R}$ $\quad \mathcal{B} = \sigma(\text{open sets}) \quad$ = Borel σ-field

$= \sigma(\text{open intervals}) = \sigma(\text{closed sets})$

$= \sigma$ (closed intervals)

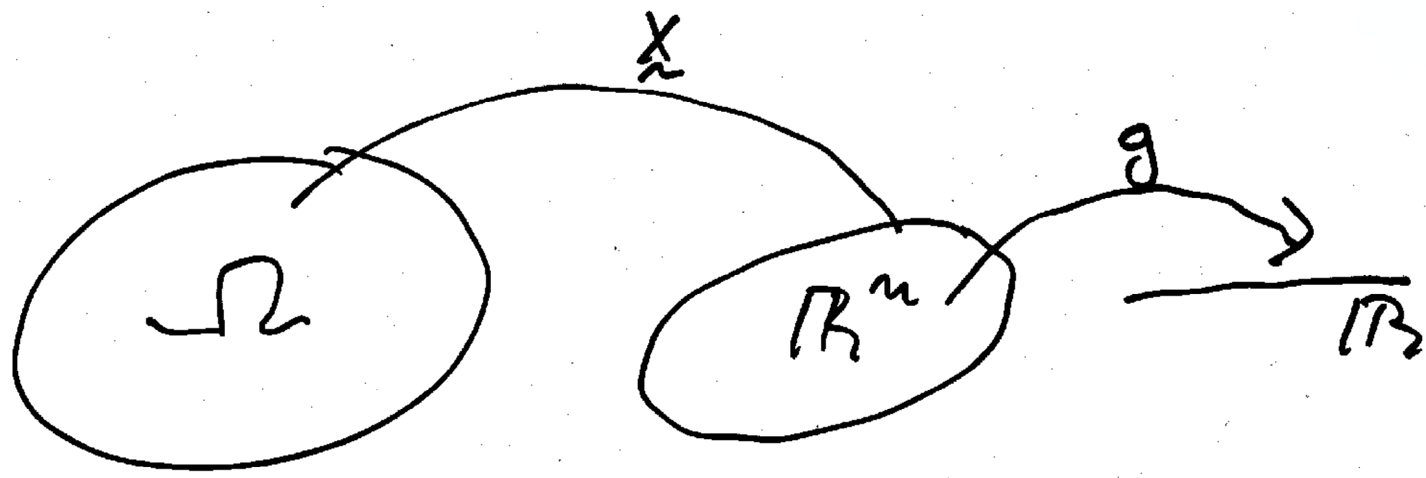
$= \sigma(\{(a, b]\}) = \sigma(\{(-\infty, b]\})$ etc...

Now let $g : \mathbb{R} \to \mathbb{R}$ be cts.
Then g^{-1}(open set) is open which
means $g^{-1}(\mathcal{B}) \subset \mathcal{B} \Rightarrow g$ is
measurable wrt $\mathcal{B}$ + $\mathcal{B}$ (also true
$g : \mathbb{R}^n \to \mathbb{R}^m$).

Look at all the cts f'ns $\mathbb{R} \to \mathbb{R}$ and
limits of convergent sequences of
cts f'ns. These are the <u>Baire</u> f'ns
+ they = set of $\underbrace{\text{mble}}_{\text{wrt } \mathcal{B}}$ f'ns

(also true from $\mathbb{R}^n \to \mathbb{R}$)

<u>Proposition</u> Let $\underset{\sim}{X} : \Omega \to \mathbb{R}^n$ be
a r vec + $g : \mathbb{R}^n \to \mathbb{R}$ be cts.
Then $g(\underset{\sim}{X}) : \Omega \to \mathbb{R}$ is mble rv.



Take a σ-field $\mathcal{F}$. Suppose for the moment that $\mathcal{F}$ is countable. The nonempty elements of $\mathcal{F}$ are

$$A_1, A_2, \ldots$$

Let $z = (z_1, z_2, \ldots)$. $\{z\}$ is uncountable

0 or 1

Set $A^z = A_1^{z_1} A_2^{z_2} \cdots$ $\quad \begin{pmatrix} A_n^0 = A_n^c \\ A_n^1 = A_n \end{pmatrix}$

$$A_n = \bigcup_{\substack{z : z_n = 1 \\ \& A^z \neq \emptyset}} A^z$$

$\Rightarrow$ countable # of $\underbrace{A^z}_{\in \mathcal{F}}$ disjoint & not $= \emptyset$

$\Rightarrow \mathcal{F}$ is uncountable

A bit more on rv's & other matters

$(\Omega, \mathcal{F}, P) \leftarrow$ probability space

$\mathcal{F} \uparrow$ σ-field

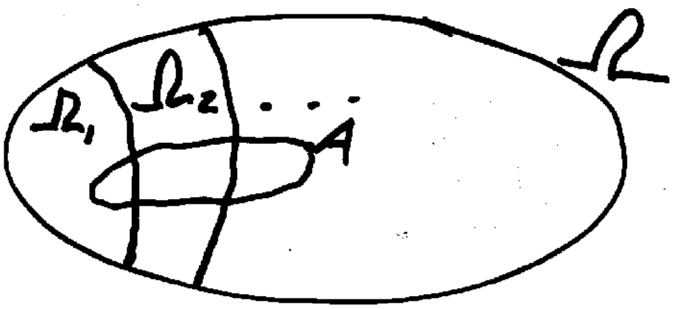
If $A \in \mathcal{F}$ & $P(A) = 0$ & $A_0 \subset A$ then we will assume $A_0 \in \mathcal{F}$. That is $\mathcal{F}$ is <u>complete</u>.

Recall P satisfies

1. $P(\Omega) = 1$ ← finite
2. $P(A) \geq 0$ ← positive
3. $P(\sum A_i) = \sum P(A_i)$ ← σ-additive

Drop 1 ⇒ <u>measure</u> $\mu(A)$. If Ω can be partitioned into sets $\Omega_1, \Omega_2, \ldots$ such that $\mu(\Omega_i) < \infty$ then μ is a σ-finite measure. In this case

$$\mu = \sum P_i$$



$$P_i(A) = \frac{\mu(A\,\Omega_i)}{\mu(\Omega_i)}$$

Note $\mu(A) = c_1 P_1(A) + c_2 P_2(A) + \cdots$. The P_i are restricted to Ω_i. In this case

$$\int X\,d\mu = c_1 \underbrace{\int X\,dP_1}_{E_1(X)} + c_2 \underbrace{\int X\,dP_2}_{E_2(X)} + \cdots$$

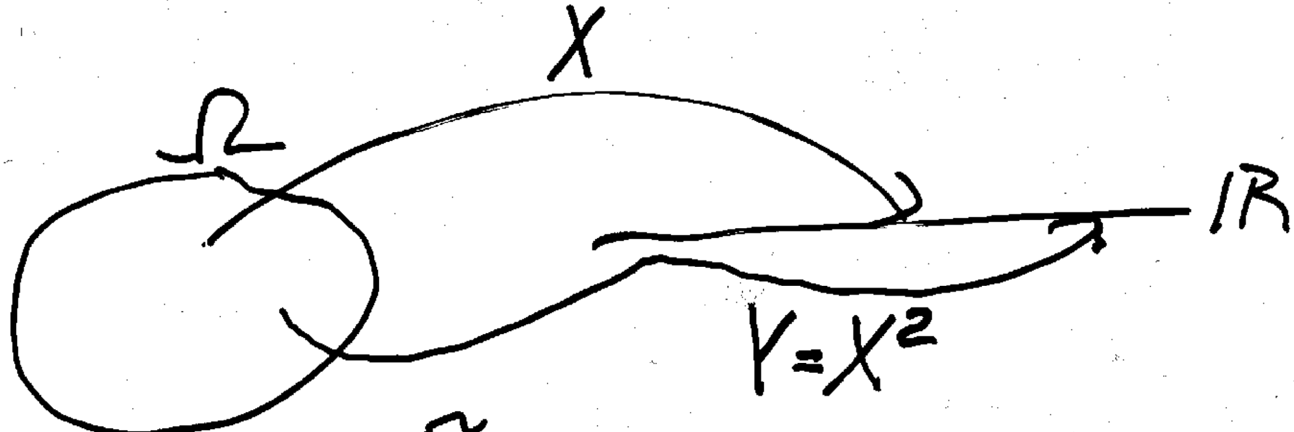
The c's are just the $\underbrace{\mu(\Omega_i)}_{c_i}$. The DCT & MCT continue to hold for these integrals.

eg Look at $\mathbb{R}$ with $\mathcal{B}$ ← Borel σ-field. Let $B \in \mathcal{B}$. Let P_i be uniform on $(i, i+1]$. Define the length of $B \in \mathcal{B}$ as

Lebesgue measure → $\lambda(B) = \sum_{i \in \mathbb{Z}} P_i(B)$

+ $\int (\ )\,d\lambda$ is the Lebesgue integral. You see it as $\int g(x)\,dx$.

Note $g : \mathbb{R} \to \mathbb{R}$ is mble wrt $\mathcal{B}$ & $\mathcal{B}$.



$\underbrace{X^{-1}(\mathcal{B})}_{\sigma\text{-field}} \subset \mathcal{F}$

$Y^{-1}(\mathcal{B}) \subset X^{-1}(\mathcal{B})$

If $Y = h(X)$ then $Y^{-1}(\mathcal{B}) \subset X^{-1}(\mathcal{B})$

Start with $X \longrightarrow$ σ-field $X^{-1}(\mathcal{B})$

σ-fields which are subsets of $X^{-1}(\mathcal{B})$ correspond to rv's which are f'ns of X.

We have defined $E(Y|X)$ as a function of X. It is convenient to define $E(Y|\underset{\substack{\uparrow \\ \sigma\text{-field}}}{\mathcal{G}})$ as that rv which is $\mathcal{G}$ & $\mathcal{B}$ mble which best predicts Y. Most of the time (for us) $\mathcal{G}$ corresponds to a rvec $\underset{\sim}{X}$ which is why we started with $E(Y|\underset{\sim}{X})$

<u>Finally</u>

For any $\underset{\substack{\uparrow \\ \text{integers}}}{n}, \underset{\uparrow}{m}$ $\exists$ k & r where

$r \in [0, |m|)$ such that

$$n = km + r$$

<u>Def'n</u> The gcd of integers n & m is an integer $k > 0$ such that $k|n$ & $k|m$ & if $l > 0$ divides n & m then $l|k$.

Remarks

1. The gcd is just the largest divisor
2. If $i, j \geq 0$ have $\gcd = 1$ then

$$\{c_1 i + c_2 j \mid \underbrace{c_1, c_2 \in \mathbb{Z}^+}_{\text{integers} \geq 0}\}$$

includes all of $\{0, 1, 2, \ldots\}$ except possibly for a finite #. Is this still true for $i_1, \ldots, i_k$ with $\gcd = 1$?

Yes.

EA $m \in \mathbb{Z}$, $n \in \mathbb{Z}^+ \Rightarrow q \in \mathbb{Z}$ & $r \in \mathbb{Z}^+$, where $0 \leq r < n$ such that

$$m = qn + r$$

If i, j are such that $\gcd = 1$ then $\exists$ integers a & b st $ai + bj = 1$

$\Rightarrow \{c_1 i + c_2 j \mid c_1, c_2 \in \mathbb{Z}^+\}$ includes all of $\mathbb{Z}^+$ except possibly a finite #.

Also true for $i_1, \ldots, i_k \geq 0$ with $\gcd = 1$

$\therefore$ $\exists$ integers $a_1, \ldots, a_k$ st $a_1 i_1 + \cdots + a_k i_k = 1$

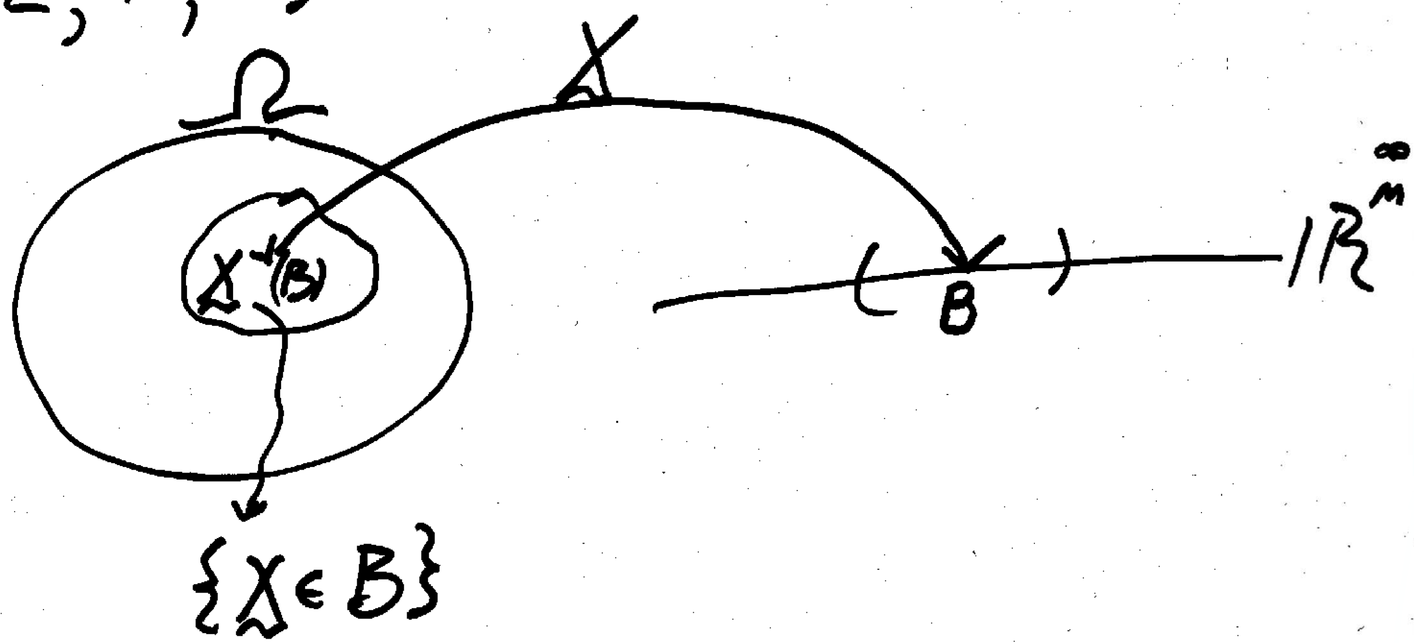
$\mathbb{R}$, $\mathcal{B}$ = Borel σ-field

$\mathbb{R}^n$, $\mathcal{B}_n$ = "

$\mathbb{R}^\infty = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix} \right\}$, $\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\}$

$\mathcal{B}_\infty$ = Borel σ-field on $\mathbb{R}^\infty$

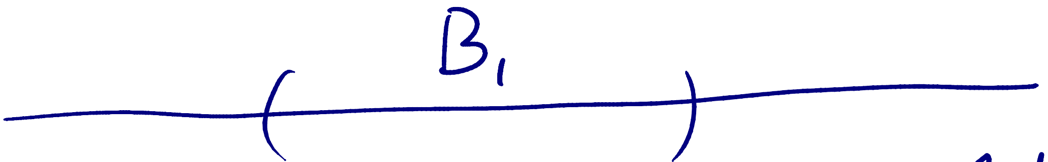
$(\Omega, \mathcal{F}, P)$



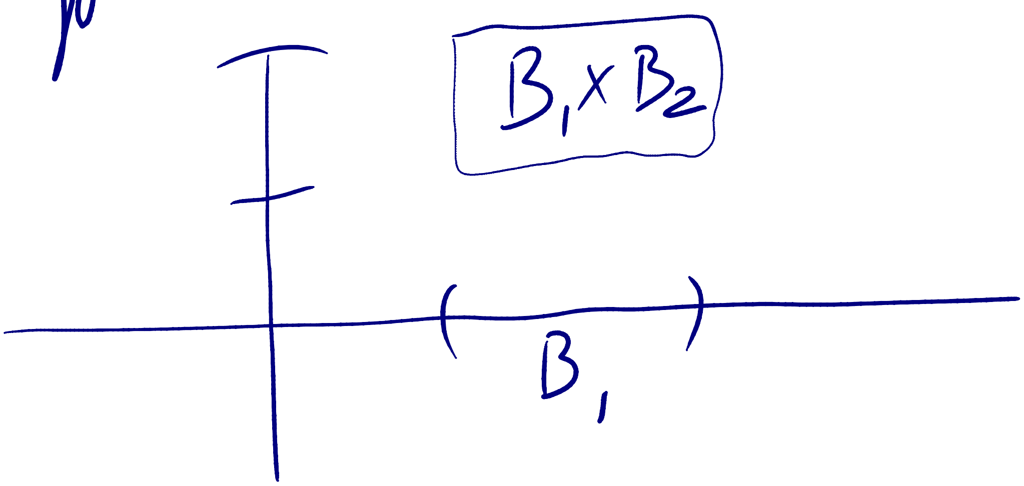
$$\underbrace{X^{-1}(\mathcal{B}_n)}_{\sigma\text{-field generated by } X \;-\; \sigma(X)} \subset \mathcal{F}$$

Let $X \in \mathbb{R}^\infty$ & suppose we want to calculate $P(X \in B)$. This can be approximated by $P(X_n \in B_n)$.

Notice that for $\mathcal{B}_1$ on $\mathbb{R}$, it is generated by sets of the type



Thus the smallest σ-field including these types of sets is $\mathcal{B}_1$. For $\mathcal{B}_2$ on $\mathbb{R}^2$ we generate it by product sets of the type



So $\mathcal{B}_2 = \sigma(\{B_1 \times B_2\})$

For $\mathbb{R}^\infty = \left\{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix}\right\}$, $\mathcal{B}_\infty = \sigma(\{B_1 \times B_2 \times \cdots\})$

If $\underset{n\times 1}{\underset{\sim}{X}}$ is a r vec (f'n from $\Omega \to \mathbb{R}^n$) then $\{ \underset{\sim}{X} \in \overset{\subset \mathbb{R}^n}{B} \}$ is an event in $\mathcal{F}$. We denote it by $X^{-1}(B)$. Formally it is $\{ \omega : \underset{\sim}{X}(\omega) \in B \}$. The collection of such events, $\{ \underset{\sim}{X}^{-1}(B) : B \in \mathcal{B}_n \}$, is a σ-field $\subset \mathcal{F}$ & is denoted by $\sigma(\underset{\sim}{X})$ or $X^{-1}(\mathcal{B})$. Notice $\underset{\sim}{X}$ induces the new probability space $(\mathbb{R}^n, \mathcal{B}_n, P_{\underset{\sim}{X}})$, where $P_{\underset{\sim}{X}}(\underset{\in \mathcal{B}_n}{B}) = P(\underbrace{\underset{\sim}{X} \in B}_{\in \mathcal{F}})$

Now take $g : \mathbb{R}^n \to \mathbb{R}$ to be cts & hence measurable wrt $\mathcal{B}_n$ & $\mathcal{B}$. The composition $g \circ \underset{\sim}{X} = g(\underset{\sim}{X})$ is of course measurable wrt $\mathcal{F}$ & $\mathcal{B}$. This only requires g to be mble wrt $\mathcal{B}_n$ & $\mathcal{B}$ (& not necessarily) cts.

Take a countably ∞ # of events $A_1, A_2, \ldots$ We denote the event $\{$an ∞ # of the A's occurs$\}$ by $\{A_n \text{ io}\}$ or just A_n io. It is also denoted by $\overline{\lim} A_n$. Notice

$$\overline{\lim} A_n = \bigcap_m \bigcup_{n \geq m} A_n$$

Borel-Cantelli

(a) $\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(A_n \text{ io}) = 0$

(b) A's ind & $\sum_{n=1}^{\infty} PA_n = \infty \Rightarrow P(A_n \text{ io}) = 1$

Proof (a)

$$\begin{aligned} P(A_n \text{ io}) &= P\Big(\bigcap_m \bigcup_{n \geq m} A_n\Big) \\ &= P\Big(\lim_{m \to \infty} \bigcup_{n \geq m} A_n\Big) \\ &= \lim_{m \to \infty} P\Big(\bigcup_{n \geq m} A_n\Big) \\ &\leq \lim_{m \to \infty} \sum_{n=m}^{\infty} P(A_n) \qquad \text{(Boole)} \\ &= 0 \qquad \text{(series converges)} \end{aligned}$$

For (b)

$$P(\{A_n \text{ io}\}^c) = P\left[\left(\bigcap_m \bigcup_{n \geq m} A_n\right)^c\right]$$

$$\overset{\text{de Morgan}}{=} P\left(\bigcup_m \bigcap_{n \geq m} A_n^c\right)$$

$$= P\left(\lim_{m \to \infty} \bigcap_{n \geq m} A_n^c\right)$$

$$= \lim_{m \to \infty} P\left(\bigcap_{n \geq m} A_n^c\right)$$

$$= \lim_{m \to \infty} \prod_{n=m}^{\infty} P(A_n^c) \qquad \text{(ind)}$$

$$= \lim_{m \to \infty} \prod_{n=m}^{\infty} [1 - P(A_n)]$$

$$\leq \lim_{m \to \infty} \prod_{n=m}^{\infty} e^{-P(A_n)} \qquad (e^{-x} \geq 1-x)$$

$$= \lim_{m \to \infty} e^{-\sum_{n=m}^{\infty} P(A_n)} = 0$$

qed

Notice that the occurrence of $\{A_n \text{ io}\}$ does not depend on any finite # of the A's. It is not a coincidence that $P(A_n \text{ io})$ is either 0 or 1 (in the independent case). This is an example of a <u>Zero-One Law</u>.

By the way, notice

eg $X = c \Rightarrow X^{-1}(\mathcal{B}) = \{\emptyset, \Omega\}$

events have prob 0 or 1

This would also be the case if $X \stackrel{as}{=} c$ (that is, $X^{-1}(\mathcal{B})$ would consist of events having prob 0 or 1.

Look at a sequence of rv's

$$\underbrace{X_1, X_2, \ldots, \underbrace{X_n, X_{n+1} \ldots}\}}_{}\ \underset{\sim}{X}$$

Look at $\underbrace{\sigma(X_n, X_{n+1}, \ldots)}_{\subset \mathcal{F}}$

$$\underbrace{\bigcap_n \sigma(X_n, X_{n+1}, \ldots)}_{\text{a } \sigma\text{-field}} = \underbrace{\underline{\underline{\text{Tail}}}\ \sigma\text{-field}}_{\text{events called } \underline{\underline{\text{Tail events}}}}$$

<u>eg</u> Let $X_1, X_2, \ldots$ be iid with mean μ + set

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

Let $A = \{\overline{X}_n \text{ converges}\}$

This is a Tail event!