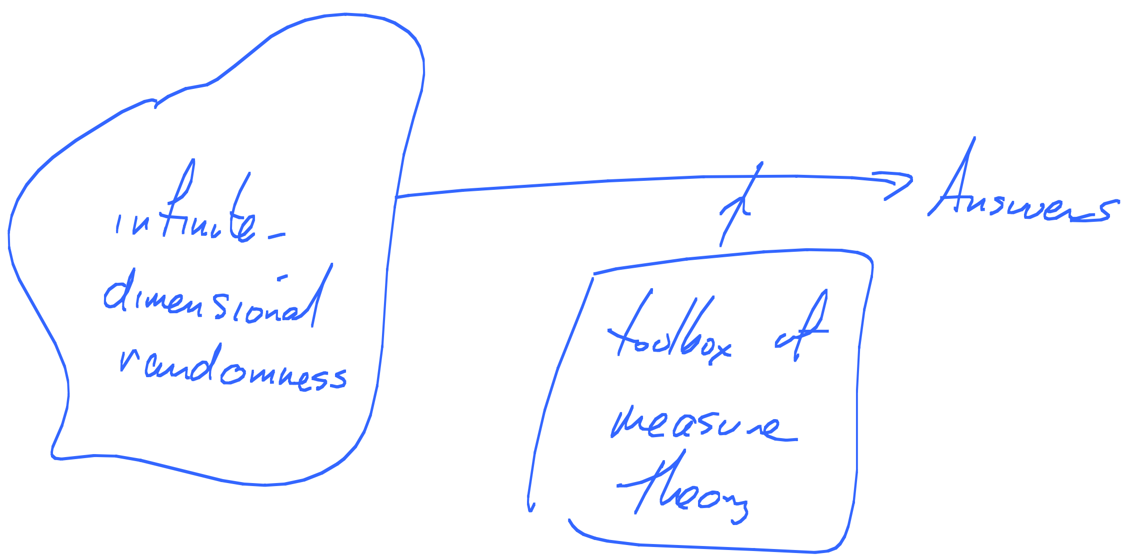
Measure Theory (& Math Review)

Note Title 1/15/2007

The grammar of probability is measure theory



Event space Ω − all possible outcomes

σ-algebra $\mathcal{F}$ of subsets of Ω

collection of "answerable questions" } important: information

$\emptyset \in \mathcal{F}$

$A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

$\{A_n\}_{n=1}^{\infty} \subset \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

$(\Omega, \mathcal{F})$ is a measurable space : Can put a measure on it

Can put many measures on it.

A probability measure $\mathbb{P}$ is a map from $\mathcal{F} \to [0,1]$ such that

a) $\mathbb{P}(\phi) = 0$

b) If $\{A_n\}_{n=1}^{\infty} \subset \mathcal{F}$ disjoint,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

c) $\mathbb{P}(\Omega) = 1$ ~ without this, just a measure

Examples:

$\Omega = \mathbb{R}$

$\mathcal{F} = \mathcal{B}(\mathbb{R})$ = smallest σ-algebra containing open subsets of $\mathbb{R}$

$$\mathbb{P}_1(A) = \chi_A(0) = \begin{cases} 1 & \text{if } 0 \in A \\ 0 & \text{else} \end{cases}$$

} Dirac at 0 δ_0

- $\mathbb{P}_2(A) = \lambda(A \cap [0,1])$ ~ Uniform $[0,1]$

  Lebesgue measure

- $\mathbb{P}_3(A) = \int_A \frac{e^{-s^2/2}}{\sqrt{2\pi}}\, ds$ ~ $\mathcal{N}(0,1)$

Ω may be complicated (e.g. $C([0,\infty))$)

$\mathcal{F}$ is often a Borel σ-algebra

$\mathbb{P}$ may be given indirectly

$(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability triple

Task is to compute

Often we want a real-valued function of randomness

$$\Omega \xrightarrow{X} \mathbb{R}$$

Definition: X is a real-valued random variable if $X^{-1}(O) \in \Omega$ for all open O

Similarly, if $(E, \mathcal{E})$ is a measurable space, a map $X: \Omega \to E$ is a random variable if $X^{-1}(S) \in \mathcal{F}$ for all $S \in \mathcal{E}$.

Expectations

A real-valued random X variable is simple if

$$X(\omega) = \sum_{k \in K} a_k \underbrace{\chi_{A_k}(\omega)}$$

$$\chi_{A_k}(\omega) = \begin{cases} 1 & \text{if } \omega \in A_k \\ 0 & \text{else} \end{cases}$$

for some finite K, finite $\{a_k\} \subset \mathbb{R}$ and $\{A_k\} \subset \mathcal{F}$. Then

$$\mathbb{E}[X] = \sum_{k \in K} a_k \mathbb{P}(A_k)$$

Note: expectation is thus linear on the set of simple functions

More generally, if X is a nonnegative random variable,

$$\mathbb{E}(X) = \sup\{ E(X') : X' \leq X \text{ pointwise}, X' \text{ simple}\}$$

If X is a random variable

$$\mathbb{E}(X) \overset{\text{def}}{=} \mathbb{E}(\underbrace{X^+}_{\max\{X,0\}}) - \mathbb{E}(\underbrace{-X^-}_{\min\{X,0\}})$$

assuming that one of these is finite

Conditional Expectations

$X \geq 0$,

$\mathcal{G}$ a sub σ-algebra of $\mathcal{F}$.

$E(X|\mathcal{G})$ is by definition a $\mathcal{G}$ measurable random variable such that

$$\mathbb{E}[\mathbb{E}(X|\mathcal{G})\chi_A] = \mathbb{E}(X\chi_A) \text{ all } A \in \mathcal{G}$$

ie a random variable ξ such that

$\xi^{-1}(A) \in \mathcal{G}$ all $A \in \mathcal{B}(\mathbb{R})$

One more thing:

In many cases, we will be able to perform calculations on a particularly "nice" (simple) set of functions, but the value of these calculations will only be apparent when we take limits of such functions

Example: We can explicitly compute the area under a piecewise constant (step) function. We can approximate a continuous function by a step function. This leads to the Riemann integral.

General mathematical background:

Banach space B

a) a vector space ~ x & y in B $\alpha, \beta \in \mathbb{R}$

$\alpha x + \beta y \in B$

b) has norm $\|\cdot\|$

triangle inequality, $\|\alpha x\| = |\alpha| \, \|x\|$,

$\|x\| = 0 \Rightarrow x = 0$

c) complete: $(x_n)_{n=0}^{\infty}$ converges.

i.e. there is an $x^* \in B$

such that $\lim_{n\to\infty} \|x_n - x^*\| = 0$

if and only if

$$\lim_{m,n \to \infty} \|x_m - x_n\| = 0.$$

Primary example: $\mathbb{R}$.

If $\lim_{m,n\to\infty} |x_m - x_n| = 0$, then there is

N such that $|x_m - x_n| < 1$ for m & $n \geq N$.

Hence $x_m \in [x_N - 1, x_N + 1]$ for all $m \geq N$.

Some extra work shows that in fact there is indeed $x^* \in \mathbb{R}$ such that

$$\lim_{n \to \infty} x_n = x^*$$

<u>Other examples</u> (which stem from the fact that $\mathbb{R}$ is a Banach space)

$L^2(\Omega)$ = collection of square integrable random variables

$$\|X\|_{L^2(\Omega)} = \mathbb{E}[|X|^2]^{1/2}$$

$L^2(\mathbb{R})$ = collection of square-integrable functions on $\mathbb{R}$

$$\|f\|_{L^2(\mathbb{R})} = \left\{ \int_{-\infty}^{\infty} |f(s)|^2 ds \right\}^{1/2}$$

$C([0,T])$ = collection of continuous functions on $[0,T]$

$$\|\varphi\|_{C([0,T])} = \sup_{0 \le t \le T} |\varphi(t)|.$$