01/29/04 Uncountable (continuous) probability spaces

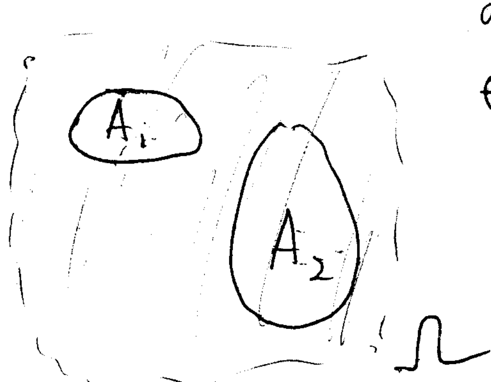
What's different from discrete, countable probability space?

- measure theory becomes technical
- doesn't make much sense to talk about Prob$(X = x)$ for random variable X with continuous state space.
  - work instead with probability densities.

Measure theory on uncountable probability space Ω

- can't always define a probability measure P on Ω such that
  - i) $P(A)$ is defined for every $A \subseteq \Omega$
  - ii) the probability measure obeys reasonable rules
    - a) $P(\Omega) = 1$
    - b) $P(\emptyset) = 0$
    - c) $P(A) \geq 0$ for every set $A \subseteq \Omega$ for which $P(A)$ is defined
    - d) $P(A^c) = P(\Omega \setminus A) = 1 - P(A)$
    - e) $P(\bigcup_{j\geq 1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j)$ if the A_j are disjoint.

      (also true for finite disjoint unions)



So to be mathematically self-consistent, for any probability measure P associated to a probability space Ω, one has to also specify a <u>σ-algebra</u> of "measurable subsets" $\mathcal{B}$.

σ-algebra $\mathcal{B}$: collection of sets, which satisfies:

i) If $A \in \mathcal{B}$, then $A^c \in \mathcal{B}$

ii) If $\{A_j\}_{j=1}^{\infty} \in \mathcal{B}$, then $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}$

iii) $\emptyset \in \mathcal{B}$

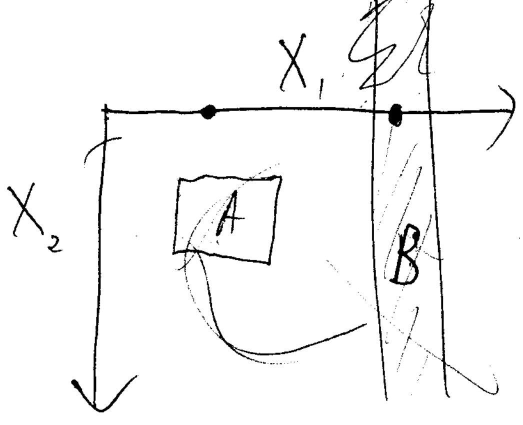
In practice:

a) rules out asking what $P(A)$ is for really crazy sets A.

b) This is a useful framework for talking about partial information, – relevant for stochastic processes

$B \in \mathcal{B}_1$

fs



Construct the σ-algebra $\mathcal{B}_1$ corresponding to events that are measurable based on info available after first day.

$A \in \mathcal{B}$ (σ-algebra for full probability space of 2 days)

but $A \notin \mathcal{B}_1$

References on Probability & Measure Theory

- Billingsley, Probability and Measure
- Sinai, Probability Theory
- Dudley, Real Analysis and Probability
- Folland, Real Analysis

Continuous random variables:

Random variables $\underline{X}$ with a continuous state space S

For example $S = \mathbb{R}$ or $S = \mathbb{R}^3$ or

$S = \mathbb{R}_{\geq 0}$ or $S = \mathbb{C}$ or $S = [0,1]$

Not meaningful to discuss $\text{Prob}(\underline{X} = x)$ for $x \in S$

Instead talk about probability of $\underline{X}$ belonging to reasonable sets.

$$\text{Prob}(\underline{X} \in B) \equiv P_{\underline{X}}(B) \quad \text{for } B \in \mathcal{B}$$

where $\mathcal{B}$ is the σ-algebra of "reasonable" subsets of S.

Reasonable subset: usually $\mathcal{B}$ to be the "Borel" sets.

- countable unions and intersections and complements of open and closed intervals.

  $S = \mathbb{R}$: $[0,1] \in \mathcal{B}$, $(0,1] \in \mathcal{B}$, $(0,1) \in \mathcal{B}$

- or sometimes $\mathcal{B}$ is taken to be the collection of Lebesgue-measurable subsets
  - include any subset of measure zero.

So a random variable X can be specified completely by defining

$$P_X(B) = \text{Prob}(X \in B) \quad \text{for any } B \in \mathcal{B}$$

with $\mathcal{B}$ a σ-algebra of (Borel) subsets of state space S.

such that P_X obeys the axioms of a probability measure.

More efficient ways of encoding this information if $S = \mathbb{R}^n$ (or $\mathbb{C}^n$)!

First for $S = \mathbb{R}$ (or some subset)

Define <u>cumulative distribution function</u>

$$F_X(x) = \text{Prob}(X \leq x) \quad \text{for } x \in \mathbb{R}$$

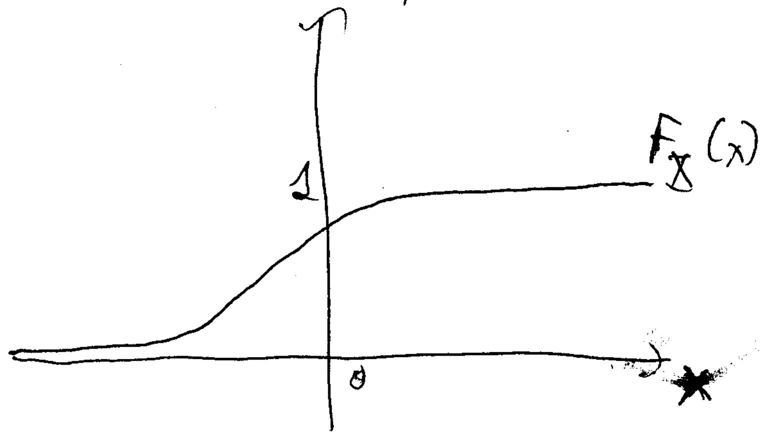
- this also completely describes the random variable

$$\text{Prob}(X \in (a,b]) = \text{Prob}(a < X \leq b) = P_X((a,b])$$
$$= F_X(b) - F_X(a)$$

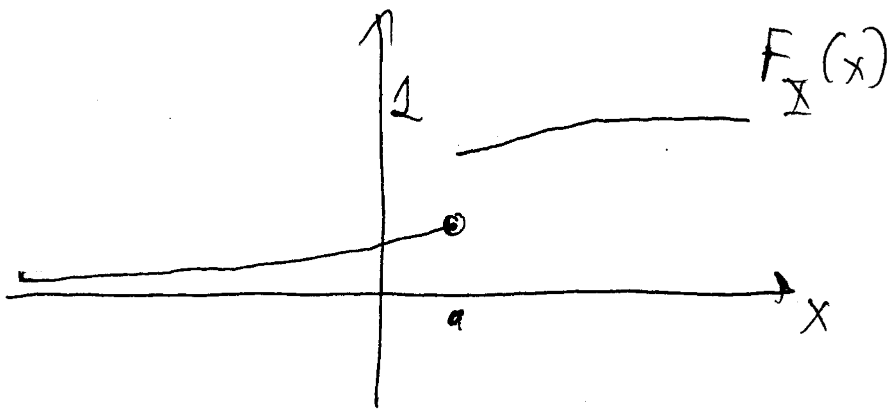
[ ]

$$\text{Prob}(X \in [a,b]) = \text{Prob}(a \leq X \leq b) = P_X([a,b])$$
$$= \text{Prob}\left(X \in \bigcap_{j=1}^{\infty} (a - 2^{-j}, b]\right)$$
$$= F_X(b) - \lim_{x \uparrow a} F_X(x)$$

If there are no points with nonzero prob.



If $x=a$ is an absorption point

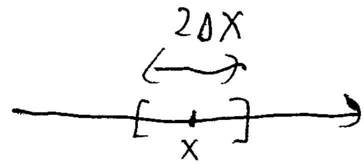


Suppose $P_{\underline{X}}$ is absolutely continuous with respect to Lebesgue measure

$\Rightarrow$ there exists a probability density function $p(x)$ such that

$$P_{\underline{X}}(B) = \text{Prob}(\underline{X} \in B) = \int_B p(x)\,dx$$

for $B \in \mathcal{B}$ (Borel sets)



Formally $\text{Prob}(|\underline{X} - x| \le \Delta x)$

$$= \int_{x-\Delta x}^{x+\Delta x} p(x')\,dx'$$

$$= p(x)\,2\Delta x + o(\Delta x) \quad \text{for } \Delta x \text{ small}$$

More generally for $S = \mathbb{R}^n$, if $D_{\vec{x}}$ is a small region containing $\vec{x} \in \mathbb{R}^n$ then

$$\cancel{\text{Prob}(\vec{x} \in D_{\vec{x}})}$$

$$\text{Prob}(\underline{X} \in D_{\vec{x}}) \simeq p(\vec{x})\,\text{Vol}(D_{\vec{x}})$$

One (sometimes) practical use for cumulative distribution function is for simulating continuous r.v.'s.

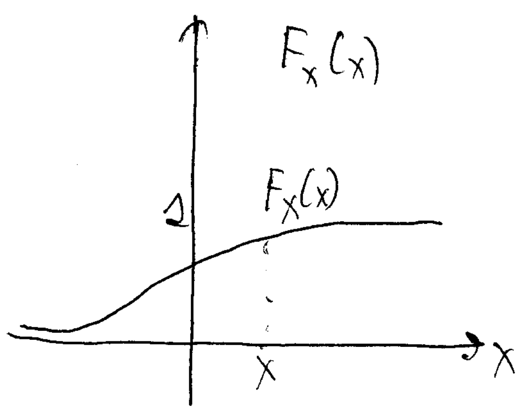
If you know the c.d.f. $F_{\mathbb{X}}(x)$ for $\mathbb{X}$, then simulate $\mathbb{X}$ by

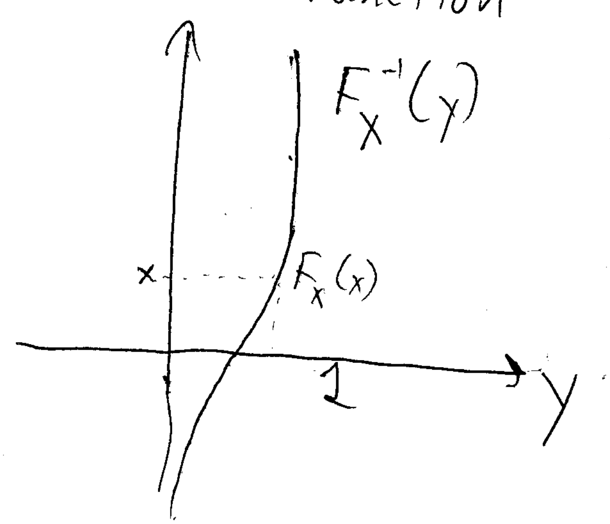
$$\mathbb{X}_{sim} = F_{\mathbb{X}}^{-1}(U)$$ where U is a uniformly distributed r.v. with $U \in [0,1]$.

Proof:

$$\text{Prob}(\mathbb{X}_{sim} \le x) = \text{Prob}(F_X^{-1}(U) \le x)$$

$$= \text{Prob}(U \le F_X(x))$$ because F_X is monotone inc. function





$$= F_X(x)$$

So $\mathbb{X}_{sim}$ has same c.d.f. as $\mathbb{X}$, so they are identically distributed r.v.'s.

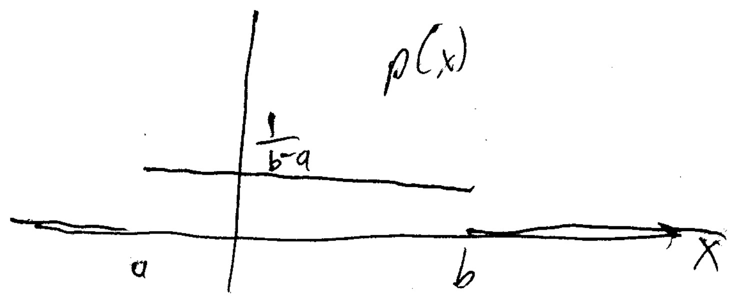
Examples:

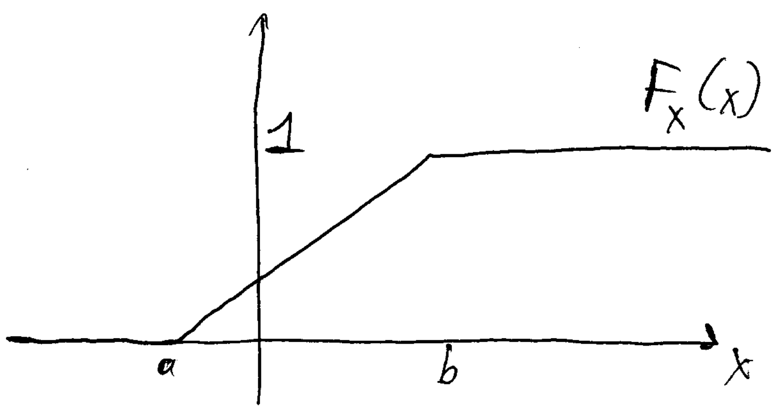
1) Uniform distribution $X \sim U[a,b]$

- X takes a value uniformly from $a \le x \le b$.

PDF: $p(x) = \frac{1}{b-a}$ for $a \le x \le b$

$= 0$ otherwise





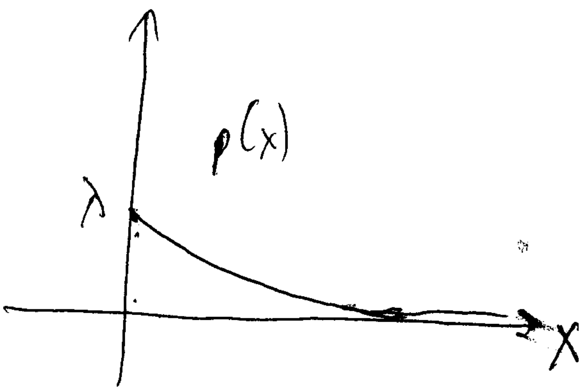
2) Exponentially distributed random variable

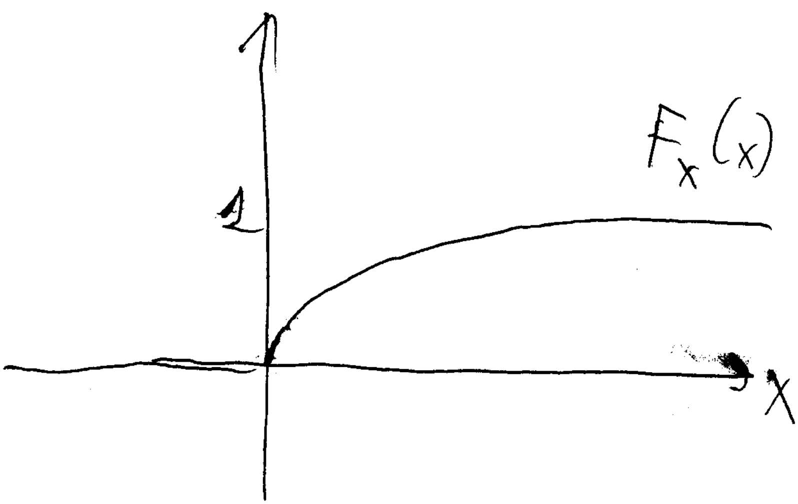
X is exponentially distributed:

PDF $\quad p(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ for some parameter $\lambda > 0$.

$p(x) = 0$ for $x < 0$

$$\langle X \rangle = \frac{1}{\lambda}$$





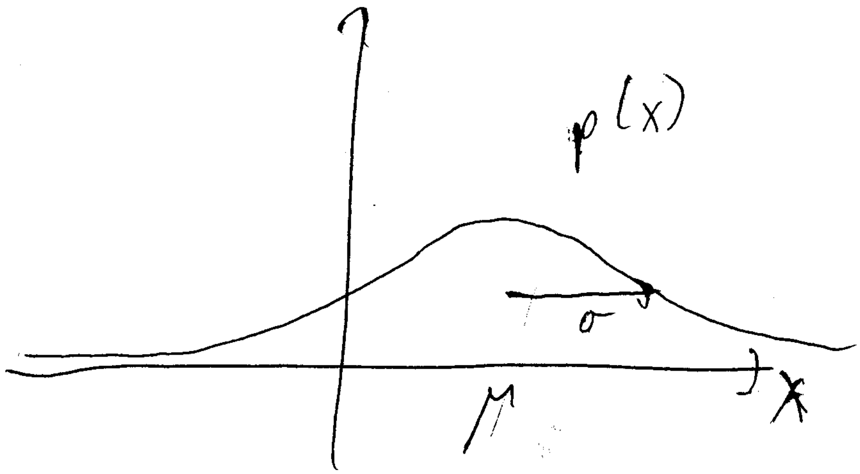
3) Gaussian distribution (normal distribution)

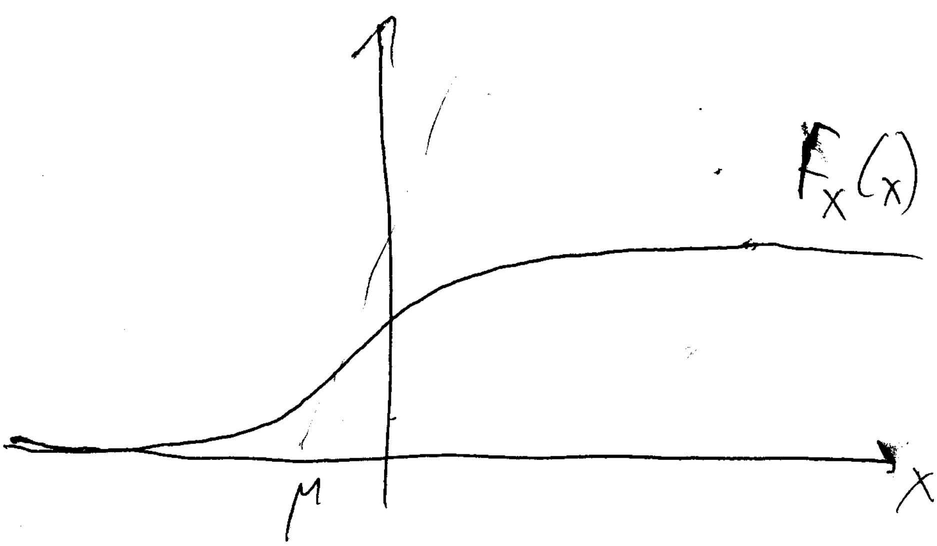
$$X \sim N(\underset{\text{mean}}{\mu}, \underset{\text{standard deviation}}{\sigma})$$

PDF $$p(x) = \frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}}$$

$$\langle X \rangle = \mu$$

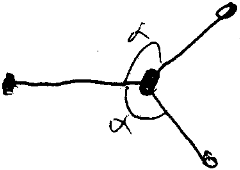
$$\langle (X-\mu)^2 \rangle = \sigma^2$$





Polymer problem:

angle α is fixed, not random.



HW 1 due 02/09 5 PM