Module 1

MODEL : High-level representation of the operations of a Real-World system.

- Discrete vs Continuous (eg. Weather)
- Stochastic ("sexier) vs Deterministic
- Dynamic vs Static

→ Models are "solved". How?

1. Analytic → EXACT Eg. Physics eq. for stone off cliff.
2. Numeric → Too tough for analytical!
3. Simulation Eg Model Weather! solve PDES NO CLOSED FORM!
   - → Add some RANDOMNESS and you are fucked. Use when analytical or numerical fail you :) VERY POWERFUL!
   - → Basically imitate some system using history (real or imagined!)

→ Simulation can give more __details__ than just an "answer".

Eg: what is waiting time in queue?

OK: Use ANAL or NUM methods

What % of customers got fed up and left?

Cannot use → Use __SIM__!

→ Use for Experimental studies!

→ Use for verification of ANA or NUM results.

→ Use for All the Good Things!

What Sucks?
- Not Easy (Sometimes)
- Time/Cost consuming.
- You don't get "The Answer"!

Gives you prob. w/confidences.

⚠ RANDOM OUTPUT!

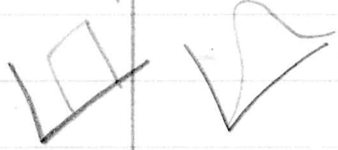
Eg. in Supply Chain?

→ Why not just use averages? Problem is __variability__. There is such a thing called "The Flaw of Averages"!

Simulation helps w/this.

Areas
percentiles!

Numbers



- What is a Continuous Function?

→ What is a limit? Really?

8 → Why not divide E[X] by n?
if its an average?

↓

$\boxed{\text{SEE}}$: josh starmer's
EV video!
STATQUEST!

→ Moment generating func [MGF]

→ $\int_{\mathbb{R}}$ ← WTF?
what are the limits?

Random Notes

→ $E[X]$ is the theoretical mean of the X
It is a parameter not a statistic!

POPULATION — SAMPLE

Params: μ, σ^2 Statistics: $\bar{x}, \hat{\sigma}^2$

This is the EV! $\longrightarrow \mu = E[X] = \sum_{\text{all } x} x\,p(x)$

$$\mu = E[g(x)] = \sum_x g(x)\,p(x)$$

$$\sigma^2 = Var[X] = E[\underbrace{(X-\mu)^2}_{}] = \sum_x (x-\mu)^2 p(x)$$

sq. dist from mean!

(do some math) $\longrightarrow E[X^2] - (E[X])^2$

$$\sigma^2 = Var[X] = E[X^2] - \mu^2$$

PMF (Discrete) Mass

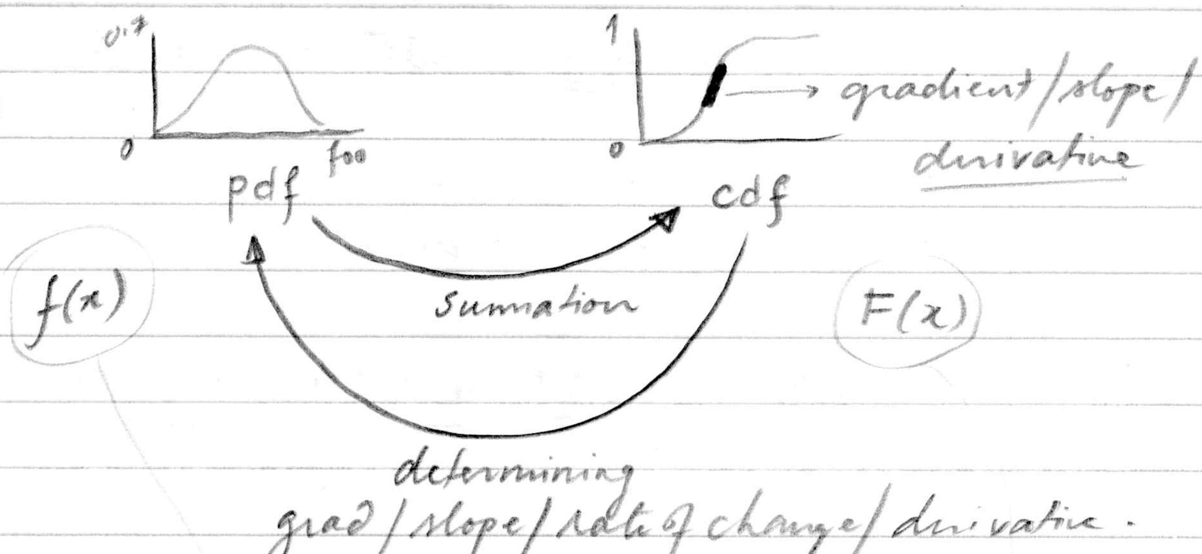
PDF (Continuous) Density

$\longrightarrow$ CDF

Cumulative Density Func.

This is really beautiful (see Zed Statistics CDF, PDF, PMF video)

(cont)

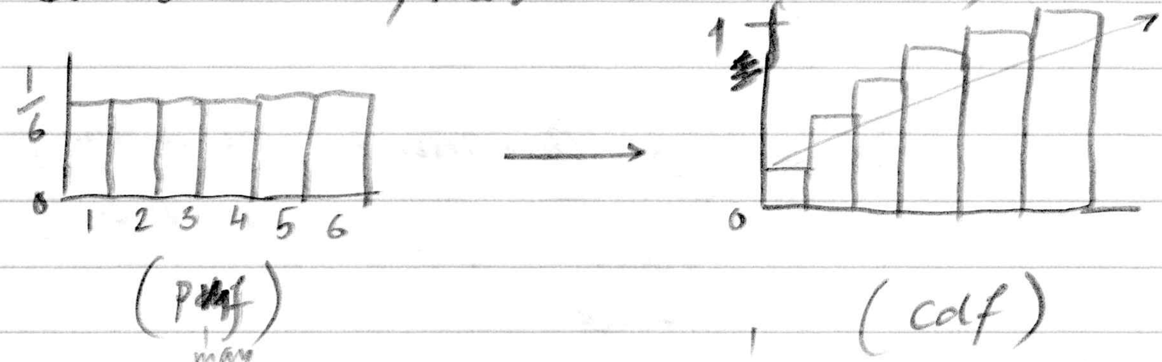


Again: pdf $\longrightarrow$ cdf : $\int_{-\infty}^{x} f(x)\,dx$
(CONT)

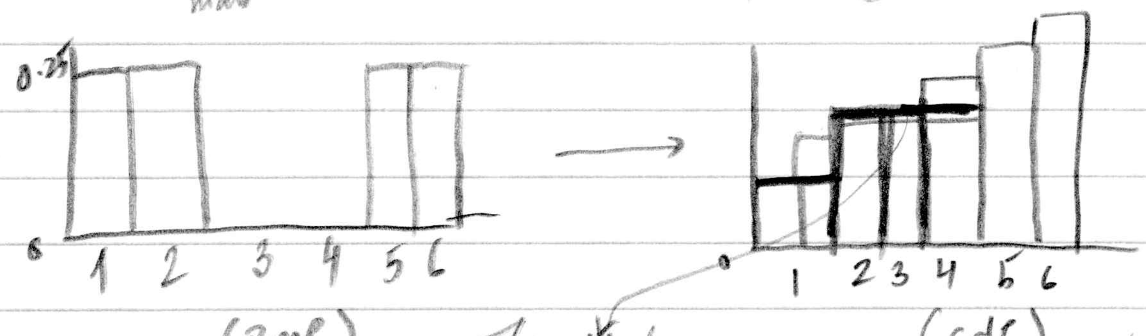
cdf $\longrightarrow$ pdf : $\frac{dF(x)}{dx} = f(x)$

In discrete, makes more (visual) sense.

fair Die



Unfair die



This flatness tells you a lot! NOW SEE CONT to make things fall into place!

→ Expected value is a
parameter, not a statistic

→ | x | -1 | $+1$ |
|---|---|---|
| $f(x)$ | 0.13 | 0.87 |

$$E[X] = \frac{(0.13)(-1)(100) + (0.87)(+1)(100)}{(100)}$$

if done 100 times

all the 100's cancel out!
regardless of the # of trials.

$$\therefore \underset{\text{(discrete)}}{E[x]} = \sum_x x f(x)$$

$$= \sum_x x P(X = x)$$

You will get this much, on average per trial
have
make
realize

StatQuest
Zed Stats
jstats
Book
Lectures.

EV Example : "I give you \$ 100 if you roll 6 You give me \$10 if not".

	1	2	3	4	5	6
$\boxed{\text{OUTCOME}}$ $x =$	-10	-10	-10	-10	-10	$+100$
$f(x) =$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

$\Rightarrow E[x] = 8.3 \Rightarrow$ you make avg \$8.3 / roll if you roll a bunch of times!

→ Average is a special case of EV where all the prob. are equal! DAMN!!

Now, $Var[x] = E\left[(X - E(x))^2\right]$

(with some math!) $= E[x^2] - \left(E[x]\right)^2$

how much does x deviat from its EV

→ $E[AX+B] = A\,E[x] + B$

→ $Var[AX+B] = A^2 \cdot Var[X]$

Ω → "Sample Space"

$\mathcal{F}$ → Curly F : Events

DISTROS

Bernoulli
Binomial
Geometric
Neg. Binomial
Poisson

DISCRETE
pmf

Uniform
Exponential
Gamma
Triangular
Beta
Normal

CONTINUOUS.
pdf

PROBABILITY

- There is a lot of stuff about
  - Sample Spaces
  - Events

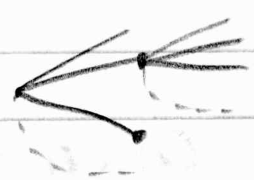
- Harvard Lecture : $\dfrac{\text{Desirable Outcomes}}{\text{TOT OUTCOMES}}$

$\therefore$ You cannot do what you have done with (eg) $1/52 \cdot 1/51$ ....

→ "Full House" Example. $\dfrac{\binom{13}{1}\binom{4}{3}\cdot\binom{12}{1}\binom{4}{2}}{\binom{52}{5}}$

777KK
444AA
666 10 10

Total outcomes ← $\binom{52}{5}$



He talks about this mult. rule.

$\boxed{\text{RANDOM VARS}}$

map $\Omega \longrightarrow \mathbb{R}$

Sample sp. — Real numbers

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

- Discrete → Finite or Countably ∞
  BERN, BINOM, POISSON, GEOM
- Continuous → Prob is zero @ every point!!
  EXP, UNIF, NORM

Discrete → PMF → CDF

Continuous → PDF → CDF

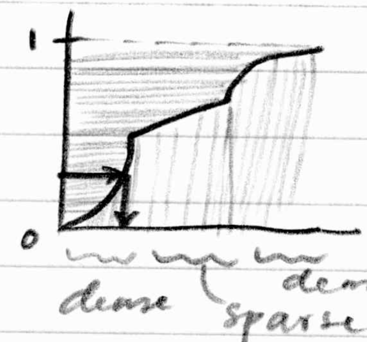
differentiation

Summation ($\int$)

→ This whole course is about <u>SIMULATING RV's</u>

$\boxed{\text{CDF}}$: Zed Stats video. It can tell us a lot about the PDF/PMF!

<u>The Inverse Transform Theorem</u>:



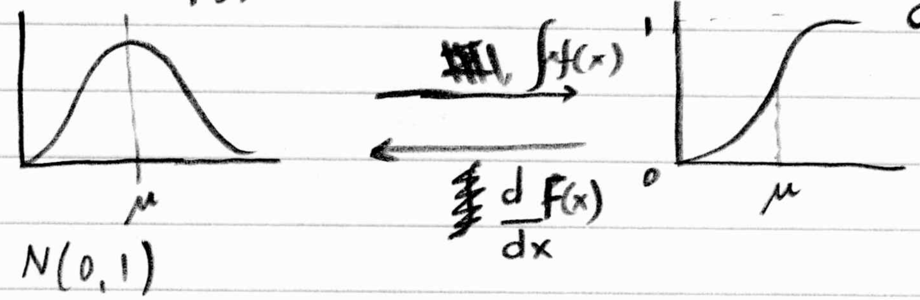
→ This can be taken to be Uniform(0,1).

→ If you have CDF, you can say things about ~~P~~/PDF without knowing about ~~the~~ PDF!

→ This is the key!

Theorem: If X is cont. w/ CDF of $F(x)$ then, $F(X) \sim Unif(0,1)$

YOU NEED TO KNOW YOUR DISTROS!



The PMF & PDF! $f(x)$

DISCRETE: $P(X=x) = f(x)$, $0 \le f(x) \le 1$ of course

It assumes a value!

CONTINUOUS: $f(x)$ is not a probability!!!

Cannot be! Think about it...

→ So: $f(x)\,dx = P(x < X < x+dx)$

→ $f(x)$ must be ≥ 0 but can be >1 !!

→ You deal w/areas. $P(X \in A) = \int_A f(x)dx$

Area

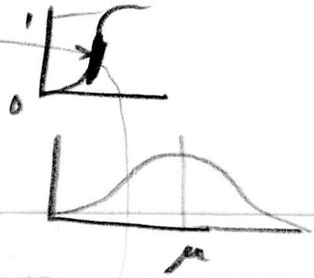
FOR BOTH, the sum of all $f(x) = 1$!

Discrete: $\sum f(x) = 1$

Cont: $\int_{\mathbb{R}} f(x)dx = 1$

$$\mu \equiv E[X] = \begin{cases} \sum x f(x) & \text{DISC} \\ \int_{\mathbb{R}} x f(x) & \text{CONT} \end{cases}$$

→ If the gradient of the CDF is steep, a lot of the distribution is around the mean in the PDF!



→ If it is ~~[illegible]~~ flatter, very little of the dist is around the mean.

Gradient! $\left(\frac{\text{rise}}{\text{run}}\right)$

BINOMIAL DIST (DISC)

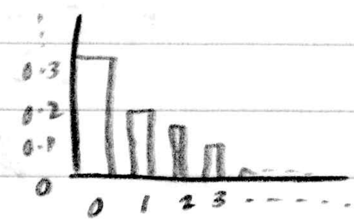
"n-number of Bernoulli trials"

ASSUME
- → p is prob. of success, $(1-p)$ fail
- → iid : One trial doesnot affect another
- ⟹ → TWO OUTCOMES ONLY PER TRIAL!!
- ⟹ → FIXED # OF TRIALS!

So, $\underbrace{p, (1-p), n}$, x | $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$

successes

pmf $f(x)$



mean $E(X) = np$

Var $Var(X) = np(1-p)$

$X \sim Bin(n, p)$

are params!

$f(x) \equiv P(X=x)$

Say ① X is $Bin(n, p)$

success!

② $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

Sai

Mama
Catherine

STOP
KILLING
YOUR
SELF

In
flowers

She was
right about
you.

You are smart.
You are capable

You need to shed some things.

Every day
Every hour, every minute, second

Time
- Fixed - increment advan Continuous
- Next - Event Time advan. Discrete!
  - FEL (update when event happens!)
    NOT BETWEEN!

M4 / PCA, dimensionality Red

Dim. Red:

$$\begin{bmatrix} \vdots \\ \vdots \end{bmatrix} \longrightarrow \begin{bmatrix} \vdots \end{bmatrix}$$

n things $\qquad$ $\underline{k}$ things

$$\mathbb{R}_n \longrightarrow \mathbb{R}_k$$

helps with cost, speed, simplicity!

→ PCA is a <u>linear</u> Dim Red technique!

You are taking pictures and turning them into 2D data (x, y) so you can plot and <u>cluster</u> and classify!

They may be high-dim, but they (math people) are trying to project the prob. into 2D space.