Continuous r.v.

a Geometric experiment is equivalent to performing many iid Bernoulli experiments with the 'stopping time' of the 1st successful Bernoulli. e.g. you keep flipping that proverbial coin until it lands Heads. It can happen on the 1st flip or on the nth.

remember ? total rv. via sum rv.

$T_n = X_1 + X_2 + \dots + X_n$

$T \sim \text{Geometric}(\ p\ ) := (1-p)^{t-1}\ p$ — PMF via p(t)

$F(t) = 1-(1-p)^t$ — CDF via F(t)

$1-F(t) = (1-p)^t$

lets overload T to also to mean time

Geometric: $T_n = \overset{0}{X_1} + \overset{0}{X_2} + \dots + \overset{1_{stop}}{X_n}$ where n≥1 — Bernoulli

- Now imagine we are conducting these iid Bernoulli experiments X_1 , X_2 ,..., X_n each every second.

1 sec — the 'stopping time' might happen at any time

1 Bernoulli | 1 Bernoulli | 1 Bernoulli | 1 Bernoulli

0 1 2 3 4

$T_n = X_1 + X_2 + \dots + X_n$ where n≥1 — time is measured in seconds — Bernoulli

- Next imagine we are able to perform n-number of iid Bernoulli experiments X_1 , X_2 ,..., X_n in a single second. Meaning, our Geometric experiment might be over in under one second.

1 sec — n Bernoulli | n Bernoulli | n Bernoulli | n Bernoulli

0 1 2 3 4

a geometric t(X) = 1/p

$E[T] = \frac{1}{p}$ experiments · $\frac{1\,sec}{n\ experiments} = \frac{1}{np}$ sec

doing so many iid Bernoullis alters our PMF $p(t) = (1-p)^{t-1}\ p$ and CDF $F(t) = 1-(1-p)^t$

if n→large and p→large we stop in a fraction of a second
if n→small and p→large we stop a little later but still in a fraction of a second

- Imagine n→large but p→small — now we have a chance to stop in a reasonable time

Let $\lambda = n\ p$ → $p := \frac{\lambda}{n}$ **reparametrization** — akin to what we did before with Hypergeometric

100,000 × 1/100,000 = 10 (reasonable #)
100,000 × 1/100,000 = 1 (reasonable #)
our experiments · prob. of success = λ

$p(t) = (1-\frac{\lambda}{n})^{t-1}\ \frac{\lambda}{n}$ and $F(t) = 1-(1-\frac{\lambda}{n})^t$

each second there is n-of experiments

PMF

- Lets make n→humongous n→∞ , BUT λ remains a reasonable constant e.g. 1,...,10 , which means that p vanishes at the same rate as n→∞ .If the experiments occur continuously, what happens to the PMF (remember PMF is a discrete function) ?

now we have n-iid Bernoullis compressed into a single second. which means that our Geometric experiment might end as soon as we press the 'start' button. Is it possible to measure the time-length of our experiment in such infinitesimally small increments? 🤔

1 sec — ∞ Bernoulli | ∞ Bernoulli | ∞ Bernoulli | ∞ Bernoulli

0 1 2 3 4

$\lim_{n\to\infty} p(t) = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{t-1}\frac{\lambda}{n} = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{t}\cdot\frac{\lambda}{n} = \lim_{n\to\infty}\frac{\lambda}{n} = 1\cdot 0 = 0$ **NO valid PMF!** (t^∞)

$\sum_{t=0\to\infty} p(t) = 0 \neq 1$ if the Σr1 , it isn't no discrete rv. 🤯 — λ=20,λ=50,λ=100

CDF

- So at the moment as n→∞ there in no valid PMF, meaning T is not a valid discrete rv. - how about CDF ?

$\lim_{n\to\infty} F(t) = \lim_{n\to\infty}\left(1-(1-\frac{\lambda}{n})^t\right) = \lim_{n\to\infty} 1 - \lim_{n\to\infty}(1-\frac{\lambda}{n})^{nt} = 1 - \lim_{n\to\infty}(1-\frac{\lambda}{n})^{nt} = 1 - \left[\lim_{n\to\infty}(1-\frac{\lambda}{n})^{n}\right]^t$

$e^{-x} = 1/e^{+x}$

$\lim_{n\to\infty} f(x)^n = \left(\lim_{n\to\infty} f(x)\right)^n$ — $(1+\frac{x}{n})^n$ → e^{+x}

$\lim_{n\to\infty}(1+\frac{x}{n})^n$

$\lim_{n\to\infty}(1+\frac{-\lambda}{n})^{nt} = \left[\lim_{n\to\infty}(1+\frac{-\lambda}{n})^{n}\right]^t = e^{-\lambda\cdot t}$

$= 1 - [\ e^{-\lambda}\ ]^t$

$= 1 - e^{-\lambda t}$

$\lim_{t\to\infty} = 1 - e^{-\lambda t}$ is this a valid CDF ?

$e := \lim_{n\to\infty}(1+\frac{1}{n})^n$

a = some reasonable const. 1...10 so if n→∞ , p→∞

let $\frac{1}{m} = \frac{a}{n} \to \frac{1}{ma} = \frac{1}{n} \to n = ma$

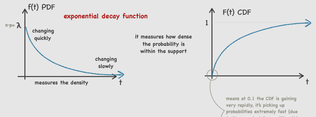
$\lim_{n\to\infty}(1+\frac{1}{n})^{na} = [\lim_{n\to\infty}(1+\frac{1}{m})^{m}]^a = e^a = e^{-\lambda\cdot t}$

PDF - Probability Density Function

PDF is a derivative of CDF — CDF continuous geometric

$f(t) := \frac{d}{dt}[\ F(t)\ ] = \frac{d}{dt}[\ 1-e^{-\lambda t}\ ] = 0 - -\lambda e^{-\lambda t} = \lambda e^{-\lambda t}$

PDF: answers the question : how dense is the probability in a certain region ? how quickly is the CDF collecting probabilities ?



it measures how dense the probability is within the support

$\int_{Supp[X]} f(x)\ dx = 1$ — the entire area = 1

$dt,0 \to 0$; $dt,\infty \to \infty$

means of 0.1 the CDF is gaining very rapidly, it's picking up probabilities extremely fast (due to the monotonicity condition this value could be very high)

Warning they are different

0.1 ≠ dt!

probability density at 1 → f(1) = 2e^{-2·1} ≈ 0.27

probability of 1 → p(1) = 0

let λ=2 — λ=2 — f(0.1) = 2e^{-2·0.1} ≈ 1.63 — p(0.1) = 0

$p(1) = P(T=1) = \int_1^1 f(t)\ dt = 0$ — very function

f(0.1) > 1 ! Possible - Yes... PDF is NOT a probability

PDF is a completely abstract metric that is good for 2 things:
1. Integrating to find the probabilities of a region via the Fund Theorem of Calc
2. Compare two points relative likelihoods

$\frac{f(0.1)}{f(1)} = \frac{1.63}{0.27} \approx 6$ — meaning the realizations near 0.1 are 6 times more likely than realizations near 1 (e.g. your geometric stopping time is to occur 6 times more likely at 0.1 than 1)

why?

$\lim_{\varepsilon\to 0}\frac{P(T\in[0.1,\ 0.1+\varepsilon])}{P(T\in[1,\ 1+\varepsilon])} = \lim_{\varepsilon\to 0}\frac{F(0.1+\varepsilon)-F(0.1)}{F(1+\varepsilon)-F(1)}$ |·$\frac{\frac{1}{\varepsilon}}{\frac{1}{\varepsilon}}$

l'Hopital's they left

$= \lim_{\varepsilon\to 0}\frac{F(0.1+\varepsilon)-F(0.1)}{\varepsilon}\Big/\frac{F(1+\varepsilon)-F(1)}{\varepsilon} = \frac{\lim_{\varepsilon\to 0}\frac{F(0.1+\varepsilon)-F(0.1)}{\varepsilon}}{\lim_{\varepsilon\to 0}\frac{F(1+\varepsilon)-F(1)}{\varepsilon}}$

PDF is a derivative of CDF
$F(t) := \int_0^t f(t)$
$F'(t) = f(t)$

$= \frac{f(0.1)}{f(1)}$

$F'(x) = \lim_{h\to 0}\frac{F(x+t)-F(x)}{h}$ — remember ext. 1!

what is the prob. of success occurring somewhere between -∞ and ∞

$P(T\in[-\infty,\infty]) = 1$ — why? $\int_{-\infty}^{\infty} f(t)\ dt = 1$ — T has to realize somehow !

$P(T\in Supp[T]) = 1 = \int_{t\in Supp[T]}^{\infty} f(t)\ dt = 1$

PDF propensity ala $\sum_{x\in Supp[X]} p(x) = 1$ — discrete rv.

$P(T=3) = p(3) = 0$ just proved
$P(T=4.9) = p(4.9) = 0$
$P(T=2.3876) = p(2.3876) = 0$
$P(T=300000) = p(300000) = 0$

??? how can a probability of everything be 0 ? paradox ?

discrete vs continuous — (it never stop - this contains infinite information)

$P(T=3) = P(T=3.00000000\dots) = 0$
$P(T=3.000) = P(\ T\in[2.99950\dots,\ 3.000499\dots]\ )$
$= P(\ T\leq 3.00049\dots\ ) - P(\ T\leq 2.99950\dots\ )$

$P(\ T\in[a,b]\ ) = \int_a^b f(t)\ dt = F(b) - F(a)$ — Fundamental Theorem of Calculus

F(3.0005) - F(2.9995) > 0 — we get some positive value out of this

proving a valid CDF

4 properties every CDF must satisfy

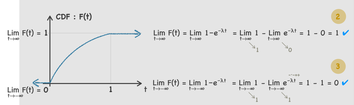
1. F(t):[0,1] — by Def probabilities must be between 0 and 1
2. $\lim_{t\to\infty} F(t) = 1$
3. $\lim_{t\to-\infty} F(t) = 0$
4. a ≤ b → F(a) ≤ F(b) ; a < b → F(a) < F(b) False — monotonically increasing, slope is ≥ 0

$1-e^{-\lambda t} \geq 0$
$1 \geq e^{-\lambda t}$ | in both sides
$0 \geq -\lambda t$
$-\lambda t \leq 0$ ✔ — $-\lambda \cdot -ve = +ve$ ✔

$1-e^{-\lambda t} \leq 1$
$-e^{-\lambda t} \leq 0$
$e^{-\lambda t} \geq 0$
$\frac{1}{e^{\lambda t}} \geq 0$ — by Def probabilities must be between 0 and 1 ; we begin at time 0, thus for geometric never ends ∞ ✔

$\sum_{t=0\to\infty} p_t = 1$

ok , so we satisfy the [0,1] 1



$\lim_{t\to\infty} F(t) = 1$

$\lim_{t\to\infty} F(t) = \lim_{t\to\infty} 1-e^{-\lambda t} = \lim_{t\to\infty} 1 - \lim_{t\to\infty} e^{-\lambda t} = 1 - 0 = 1$ ✔ 2

$\lim_{t\to 0} F(t) = 0$

$\lim_{t\to 0} F(t) = \lim_{t\to 0} 1-e^{-\lambda t} = \lim_{t\to 0} 1 - \lim_{t\to 0} e^{-\lambda t} = 1 - 1 = 0$ ✔ 3

remember calc-1 ?
→ slope is the derivative

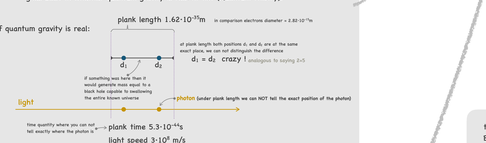
$\frac{d}{dt}[\ F(t)\ ] = \frac{d}{dt}[\ 1-e^{-\lambda t}\ ] = 0 - -\lambda e^{-\lambda t} = \frac{\lambda}{e^{\lambda t}} \geq 0$ ✔ 4 — Chain Rule

slope = 0
positive slope
negative slope

Does Continuity exist?

Is Time Discrete or Continuous? It is a big open question in Physics. We do not know. It goes back to whether quantum gravity is real or not (Quantum Theory).

if quantum gravity is real:



if gravity is quantized, then time is discrete , which means there is no such thing as Continuous r.v. 🤔

Support of Cont. r.v.

Thus squeezing infinitely many Bernoullis into a single second broke our PMF, which means that our Geometric experiment is NOT a discrete r.v. . However, it does have a CDF. If you have a CDF , you have a r.v. . It is not like any random variable we have seen before, it is a new kind of a r.v. . T is a **Continuous r.v.**

it is unworkably close to 0

$Supp[T] = (0,\infty) = [0,\infty) \neq [0,\infty]$

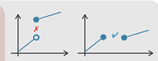
$|\ Supp[T]\ | = |\mathbb{R}|$ > $|\mathbb{N}|$ — size of a discrete r.v.

$\mathfrak{c}$ > $\aleph_0$ — uncountable infinity > countable infinity

self-note:
'what if' - the faster we go, the slower time goes & once the speed of light is reached - time stops (discrete stop).
what if the mass instead of increasing decreases and completely converts into massless wave-energy once the speed of light is breached?

Properties of Continuous r.v.

1. | Supp[X] | = | ℝ | — continuum
2. Has valid CDF F(x) no jumps but gaps allowed
3. PMF does not exist
4. PDF exists f(x)≥0 and $\int_{Supp[X]} f(x)\ dx = 1$ — the entire area = 1



Def. of Identical Distribution for Cont. r.v.s

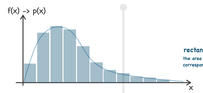
X_1 , X_2 Cont. rv. and $X_1 \overset{d}{=} X_2$ if $F_1(x) = F_2(x)$ → PDFs are the same
if $F_1(x) = F_2(x)$ → CDFs are the same — always exist
→ no PMFs

PDF - probability densities

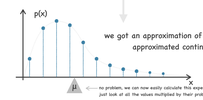
the problem is that this is a PDF (these are probability densities not probabilities)

E[X] = ?

lets discretize this cont. rv. by taking 'Midpoint Rule' approximations 'backwards'



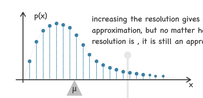
rectangle areas are probabilities
the area of each rectangle (area=prob.) corresponds to the amount of possible realizations



we got an approximation of a PDF in form of a PMF approximated continuous r.v. is a discrete r.v.

no problem, we can now easily calculate this expectation-μ just look at all the values multiplied by their probabilities and add them up

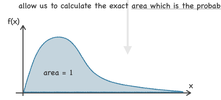
$E[X] = \sum_{x\in Approx Supp[X]} x\ p(x)$



increasing the resolution gives us a better approximation, but no matter how good the resolution is , it is still an approximation

remember? Riemann sums

allowing the separation between each value to approach 0 will allow us to calculate the exact area which is the probability



$E[X] = \int_{x\in Supp[X]} x\ f(x)\ dx$

by 'Law proof' using Lebesgue Integral

$E[\ g(x)\ ] = \sum_{x\in Supp[X]} g(x)\ p(x)$

similar

$E[\ g(x)\ ] = \int_{x\in Supp[X]} g(x)\ f(x)\ dx$

$Var[X] = E[\ (X-\mu)^2\ ]$ — $g(x)$

$Var[X] = \int_{x\in Supp[X]} (x-\mu)^2\ f(x)\ dx$

same as before

$E[aX+c] = aE[X]+c$
$Var[aX+c] = a^2\ \sigma^2$
$SE[aX+c] = |a|\ \sigma$
$E[\Sigma X_i] = \Sigma\ E[X_i] = n\ E[X]$ — if identically distributed
$Var[\Sigma X_i] = \Sigma\ Var[X_i] = n\ Var[X]$ — if independent, if iid

$Var[Z] = E[Z^2] - \mu^2$

Exponential r.v.

thus we took a Geometric and drove an infinite amount of iid Bernoullis into each second. It turns out that by doing that we arrived at a very famous Cont. rv. called **Exponential rv.**

λ ∈ ℝ
λ = n·p , i.e. #-of#_trials · prob.

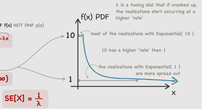
$X \sim Exponential(\ \lambda\ ) := \lambda e^{-\lambda x}$

$Supp[X] = [0,\infty)$ — 'rate'

Parameter Space λ ∈ (0,∞)

$E[X] = \frac{1}{\lambda}$ $Var[X] = \frac{1}{\lambda^2}$ $SE[X] = \frac{1}{\lambda}$

PDF f(x) NOT PMF p(x)



λ is a tuning dial that if cranked up, the realizations start occurring at a higher 'rate'
most of the realizations with Exponential(10)
10 has a higher 'rate' than 1
the realizations with Exponential(1) are more spread out

Geometric is called 'memoryless' ... Is Exponential also 'memoryless' ?

Exponential has a **memorylessness** property

$P(\ X>a+b\ |\ X>b\ ) = P(\ X>a\ )$ ✔

$P(\ X>a\ ) = e^{-\lambda a}$

$\frac{P(\ X>a+b\ \&\ X>b\ )}{P(\ X>b\ )} = \frac{P(\ X>a+b\ )}{P(\ X>b\ )} = \frac{e^{-\lambda(a+b)}}{e^{-\lambda b}} = \frac{e^{-\lambda a}\cdot e^{-\lambda b}}{e^{-\lambda b}} = e^{-\lambda a}$

$P(\ X\leq x\ ) = F(x) = 1-e^{-\lambda x}$ — CDF — use CDF to measure probability

$P(\ X>x\ ) = 1-F(x) = e^{-\lambda x}$

$E[X] = \int_0^{\infty} x\ f(x)\ dx = \int_0^{\infty} x\ \lambda e^{-\lambda x}\ dx$ | pull the const. through

$= \lambda \int_0^{\infty} x\ e^{-\lambda x}\ dx$

integration by parts: $\int u\ dv = uv - \int v\ du$

let u=x , dv=e^{-λx} dx → du=dx → v=$\frac{1}{-\lambda}e^{-\lambda x}$
→ $\int v\ du = \int \frac{1}{-\lambda}\ e^{-\lambda x}\ dx = \frac{1}{-\lambda^2}e^{-\lambda x}$

$= \lambda\ [\ -\frac{1}{\lambda}\ x\ e^{-\lambda x} - \frac{1}{\lambda^2}\ e^{-\lambda x}\ ]_0^{\infty}$

$= -(\ (\ \lim_{x\to\infty} xe^{-\lambda x} + \lim_{x\to\infty}\frac{1}{\lambda}\ e^{-\lambda x}\ ) - (\ 0\cdot e^{-\lambda\cdot 0} + \frac{1}{\lambda}e^{-\lambda\cdot 0}\ )\ )$

$= -(\ (\ 0+0\ ) - (0+\frac{1}{\lambda})\ )$

so makes sense !

$E[X] = \frac{1}{\lambda}$

also remember our Geometric journey?
$X \sim Geometric(\ p\ )$
$E[X] = \frac{1}{p}$ if n=1 → p=λ.
$\frac{1}{\lambda}$
$\frac{1}{np}$
$\frac{1}{\lambda}$

alternative way

$\int u\ v'\ dx = u\ v - \int u'\ v\ dx$

u = x , v' = e^{-λx}
u' = 1 , v = $\frac{1}{-\lambda}e^{-\lambda x}$

$= \lambda\ [\ x\ \frac{1}{-\lambda}e^{-\lambda x} - \int_0^{\infty}\frac{1}{-\lambda}e^{-\lambda x}\ dx\ ]$

$= \lambda\ [\ x\ \frac{1}{-\lambda}e^{-\lambda x} - \frac{1}{-\lambda^2}\ \frac{1}{e^{\lambda x}}\ ]_0^{\infty}$

	single stop	multi stop
discrete	Geometric	NegativeBinomial
continuous	Exponential	Erlang

Var[X] is derived also using the integration by parts - hw

Standard Normal r.v.

- find σ²

$Var[X] = E[X^2] - \mu^2 = \int_{x\in Supp[X]} (x-\mu)^2\ f(x)\ dx$ — σ²>0

$= \int_{-\infty}^{\infty} x^2\ \frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\ dx$

= integration by parts ...

$= 1 = Var[X] = \sigma^2$

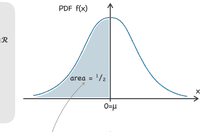
$SE[X] = \sqrt{\sigma^2} = \sqrt{1} = 1$

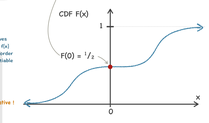
$Z \sim N(\ 0,1\ ) := \frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}$ — PDF f(x)

'Standard Normal r.v.' aka 'Bell Curve' aka 'Gaussian r.v.'

is this a PDF ? → continuous r.v.
1. f(x) ≥ 0 yes b/c $e^{-\frac{x^2}{2}} = +$ x∈ℝ
2. $\int_{x=-\infty\to\infty} f(x)\ dx = 1$ ← proof

$SE[X] = \sigma = 1$ $Var[Z] = \sigma^2 = 1$ $E[Z] = \mu = 0$ $Supp[Z] = \mathbb{R}$





- find μ

$E[X] = \int_{x=-\infty}^{\infty} x\ f(x)\ dx$ — PDF, Supp[X] = ℝ

$= \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{\infty} x\ e^{-\frac{x^2}{2}}\ dx$ | pull the const. through

let u=$\frac{x^2}{2}$, du=x dx

$= \frac{1}{\sqrt{2\pi}}\ [-e^{-\frac{x^2}{2}}]_{x=-\infty}^{x=\infty}$

$\int e^{-u}\ du = -e^{-u} = -e^{-\frac{x^2}{2}}$

$= \frac{1}{\sqrt{2\pi}}\ [\ \lim_{x\to\infty} -\frac{1}{e^{\frac{x^2}{2}}} - \lim_{x\to-\infty} -\frac{1}{e^{\frac{x^2}{2}}}\ ] = 0 = \mu$

$\frac{1}{-\infty^2} = 0$; $\frac{1}{\infty^2} = 0$

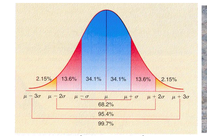
- find CDF

$F(x) = \int f(x)\ dx + C$

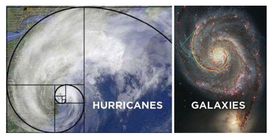
$= \int \frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\ dx + C$

It has been proven that it is impossible to find this antiderivative ! → NO CLOSED FORM
go! to use a computer to do this

Risch Algorithm gives you the rules what f(x) has to look like in order to be anti-differentiable



'68-95-997 rule'



$\int_{-1\sigma}^{1\sigma}\frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\ dx$

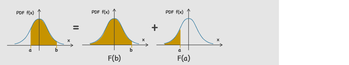
$\int_{-2\sigma}^{2\sigma}\frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\ dx$

$\int_{-3\sigma}^{3\sigma}\frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\ dx$

	PMF	PDF	CDF	E[X]	Var[X]	| Supp[X] |	Quantile[X]
Discrete r.v.	p(x)=[0,1] $\sum_{x\in Supp[X]} p(x)=1$	no PDF	Yes — not always in closed form	$\sum_{x\in Supp[X]} x\ p(x)$	$\sum_{x\in Supp[X]} (x-\mu)^2\ p(x)$	≤ |𝒩|	min {x: F(x) ≥ p}
Continuous r.v.	no PMF	f(x) ≥ 0 ; $\int_{x\in Supp[X]} f(x)\ dx = 1$ 'integrates to unity'	Yes — b/c CDF is monotonically increasing ; f(x) = ^d/_dx [CDF]	$\int_{x\in Supp[X]} x\ f(x)\ dx$	$\int_{x\in Supp[X]} (x-\mu)^2\ f(x)\ dx$	= |ℝ| > |𝒩|	x s.t. F(x) = p

68-95-997 rule

$P(\ Z\in[a,b]\ ) = F(b) - F(a)$



$P(\ Z\in[-1,1]\ ) = F(1) - F(-1) \approx .68$
$P(\ Z\in[-2,2]\ ) = F(2) - F(-2) \approx .95$
$P(\ Z\in[-3,3]\ ) = F(3) - F(-3) \approx .997$

'68-95-997 rule' aka '36 rule' aka 'empirical rule'

the amount of probability between -3SE and 3SE is .997 or 99.7%

let X~Uniform(λ) , Y=2X how is this distributed ?

a clever trick - look at the CDF of Y

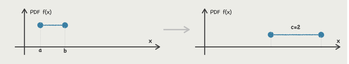
$F_Y(x) = P(\ Y\leq x\ ) = P(\ 2X\leq x\ ) = P(\ X\leq\frac{x}{2}\ ) = F_X(\frac{x}{2})$ — now we can use the CDF of X

$= 1 - e^{-\lambda\frac{x}{2}}$
$= 1 - e^{-\frac{\lambda}{2}x}$
$= 1 - e^{-\lambda' x}$ — if λ' = $\frac{\lambda}{2}$
$= F_Y(x)$

Exponential CDF $F(x) = 1-e^{-\lambda x}$

Y~Exponential($\frac{\lambda}{2}$)
Y~Exponential(λ') — s.t.

let X~Uniform(a,b) , Y=cX+d how is this distributed ?



$F_Y(x) = P(\ Y\leq x\ ) = P(\ cX\leq x-d\ ) = P(\ X\leq\frac{x-d}{c}\ ) = F_X(\frac{x-d}{c})$

P(cX≤d≤x)

$= \frac{\frac{x-d}{c}-a}{b-a} = \frac{x-d-ac}{cb-ca} = \frac{x-(ca+d)}{cb-ca} = \frac{x-(bc+d)}{cb-ca}$

let a' = ac+d , b' = bc+d

Uniform CDF $F(x) = \frac{x-a'}{b'-a'}$

Y~Uniform(ac+d , bc+d)
Y~Uniform(a',b') — s.t.

General Normal r.v.

$Z \sim N(\ 0,1\ )$ Let X=μ+σZ how is this distributed ?

$F_X(x) = P(\ X\leq x\ ) = P(\ \mu+\sigma Z\leq x\ ) = P(\ \sigma Z\leq x-\mu\ ) = P(\ Z\leq\frac{x-\mu}{\sigma}\ ) = F_Z(\frac{x-\mu}{\sigma})$ ✗

unknown since $F_Z(x)$ ('bell curve' CDF) is unknown

but...

$F_X(x) = {}^d/_{dx}\ [\ F_X(x)\ ] = {}^d/_{dx}\ [\ F_Z(\frac{x-\mu}{\sigma})\ ]$

let $u=\frac{x-\mu}{\sigma}$, $\frac{du}{dx}=\frac{1}{\sigma}$, dx=σdu

$= \frac{d}{du}\frac{du}{dx}\ [\ F_Z(u)\ ]$
$= \frac{d}{du}\frac{1}{\sigma}\ [\ F_Z(u)\ ]$

we know this $\frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}$

$= \frac{1}{\sigma}\ f_Z(u)$
$= \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\ e^{-\frac{u^2}{2}}$
$= \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}}$
$= \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

E[X] Var[X]

$Z \sim N(\ \mu,\sigma^2\ ) = F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ 'General Normal Distribution'

Param Space: μ ∈ ℝ — we can center it anywhere ; σ² ∈ (0,∞) — squeeze > stretch/flatten

μ=0 ; σ=1 — special case — 'Standard Normal Distribution'

$Supp[X] = \mathbb{R}$

we can easily find E[X] and SE[X] by plugging in for Xsp=ℝ

$E[X] = E[\ \mu+\sigma Z\ ] = \mu$
$SE[X] = SE[\ \mu+\sigma Z\ ] = \sigma$

gives a very flexible 'bell curve' (squeeze/stretchable and centered dynamically)

Uniform Cont. r.v.

discrete uniform r.v.
$X \sim Uniform(\ 1,7,28\ )$

continuous Uniform r.v. $X \sim Uniform(\ a,b\ ) := \frac{1}{b-a}$ — PDF f(x)

$Supp[X] = [a,b]$ — any value between a & b

Parameter Space a ∈ ℝ, b ∈ ℝ, a<b

$Var[X] = \frac{(b-a)^2}{12}$ $E[X] = \frac{b+a}{2}$

CDF $P(\ X\leq x\ ) = F(x) = \frac{x-a}{b-a}$

Look at the relationship between the PDF & the CDF (function and its derivative)
On the CDF graph between a & b the slope is rising at the constant rate which is unexplanatorily depicted on the graph of the PDF.

- find μ

$E[X] = \int_{x\in Supp[X]}^{b} x\ f(x)\ dx$ — PDF, Supp[X] = [a,b]

$= \int_a^b x\ \frac{1}{b-a}\ dx$ | pull the const. through

$= \frac{1}{b-a}\int_a^b x\ dx$

$= \frac{1}{b-a}\ [\ \frac{x^2}{2}\ ]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$

$\frac{1}{b-a}\ (\frac{b^2}{2} - \frac{a^2}{2}) = \frac{b^2}{2(b-a)} - \frac{a^2}{2(b-a)} =$

- find CDF

$F(x) = \int f(x)\ dx + C$

$= \int \frac{1}{b-a}\ dx + C$

$= \int 1\ dx + C$

$= \frac{1}{b-a}\ x + C$

$= \frac{x}{b-a} + C$

BUT 'C' can NOT just be anything b/c in our scenario the CDF has strict limits for what it can be

Find C. Lets use what we know

$F(a) = 0$ $\frac{a}{b-a} + C = 0$
$F(b) = 1$ $C = -\frac{a}{b-a}$

$= \frac{x}{b-a} + -\frac{a}{b-a}$

$= \frac{x-a}{b-a}$

- find σ²

expected long run average squared error

$\sigma^2 = Var[X] = \int_{x\in Supp[X]}^{b} (x-\mu)^2\ f(x)\ dx = E[X^2] - \mu^2$

$= \int_a^b x^2\ \frac{1}{b-a}\ dx - (\frac{b+a}{2})^2$

$= \frac{1}{b-a}\int_a^b x^2\ dx - (\frac{b+a}{2})^2$

$= \frac{1}{b-a}\ [\ \frac{x^3}{3}\ ]_a^b - (\frac{b+a}{2})^2$

$= \frac{x^3}{3(b-a)}\Big|_a^b - (\frac{b+a}{2})^2$

$= \frac{b^3}{3(b-a)} - \frac{a^3}{3(b-a)} - (\frac{b+a}{2})^2$

$= \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4}$

$= \frac{b^3+ab^2+a^2b-a^3}{3} - \frac{(b+a)^2}{4}$

$= \frac{(4b^3+4ab^2+4a^2) - (3b^2+6ab+3a^2)}{12}$

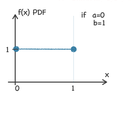
$= \frac{b^3-3ab+a^3}{12}$

$b-a$ b^3-a^3 $-(\ b^3-ab^2\ )$ $-ab^2+a^3$ $-(\ a^2b-a^3\ )$ a^3b-a^2b 0

$= \frac{(b-a)^3}{12}$ → $SE[X] = \sqrt{\sigma^2} = \sqrt{\frac{(b-a)^2}{12}} = \frac{b-a}{\sqrt{12}}$

$SE[X] = \frac{b-a}{\sqrt{12}}$

$X \sim Uniform(\ 0,1\ ) = 1$ — PDF



Standard Uniform

pick a number between 0 and 1 at random - it is the algorithmic foundation for all randomness inside computers , every programing language has a way of pulling a realization out of this.
Java - math.random
C++ rand()