10

recap…

$X \sim \text{Bernoulli}(p) := p^x (1-p)^{1-x}$

$X \sim \text{Binomial}(n,p) := \binom{n}{x} p^x (1-p)^{n-x}$

$X \sim \text{Hypergeometric}(n,K,N) := \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$

$X \sim \text{Hypergeometric}(n,p,N) := \frac{\binom{pN}{x}\binom{(1-p)N}{n-x}}{\binom{N}{n}}$

$\lim_{N\to\infty} X \sim \text{Hypergeometric}(n,p,N) = X \sim \text{Binomial}(n,p)$ conceptually…

Binomial distribution Proof

property of PMF: $\sum_{x\in Supp[X]} p(x) = 1$

if the Binomial it truly a r.v. we must prove = 1 how ?

verify $\sum_{x\in[0,\dots,n]} \binom{n}{x} p^x(1-p)^{n-x} = 1$

Binomial $Supp[X] = \{0,1,\dots,n\}$

recall the **Binomial Theorem** $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$ let $a=p$, $b=1-p$, $i=x$

reason why we call this distribution 'binomial'

plug in

$(p+(1-p))^n = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}$

$(p-p+1)^n =$

$(1)^n =$

$1 =$ QED

what are "iid" r.v.'s?

X_1 & X_2 are independent r.v. if :

$P(X_1=x_1 \mid X_2=x_2) = P(X_1=x_1)$

$P(X_2=x_2 \mid X_1=x_1) = P(X_2=x_2)$

$P(X_1=x_1, X_2=x_2) = P(X_1=x_1)\, P(X_2=x_2)$ — and ∩

JMF Joint Mass Function

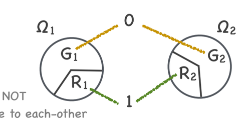
$\forall x_1 \in Supp[X_1]$ $\forall x_2 \in Supp[X_2]$

remember test for independence ?
$P(A|B) = P(A)$
$P(B|A) = P(B)$
Multiplication Rule
$P(B\cap A) = P(B)\cdot P(A)$

X_1 & X_2 are **independent & identically distributed – iid**
if X_1 and X_2 are independent & $X_1 \overset{d}{=} X_2$ (equal in distribution) denoted $X_1\ X_2 \overset{iid}{\sim}$

(statistics) patients in a medical trial are considered iid

the only way to know independence is to know Ω and the process or be given the JMF (Joint Mass Function)



spinners do NOT communicate to each-other

$X_1 \sim$ Bernoulli (1/3)
$X_2 \sim$ Bernoulli (1/3)

$X_1 \overset{d}{=} X_2$? – Yes

are X_1, X_2 independent ? – Yes, since

$P(X_1=1 \mid X_2=1) = P(R_1|R_2) = P(R_1) = P(X_1=1)$
$P(X_1=1 \mid X_2=0) = P(R_1|G_2) = P(R_1) = P(X_1=1)$
$P(X_1=0 \mid X_2=1) = P(G_1|R_2) = P(G_1) = P(X_1=0)$
$P(X_1=0 \mid X_2=0) = P(G_1|G_2) = P(G_1) = P(X_1=0)$

thus X_1, X_2 are $\overset{d}{=}$ & independent → iid (independent & identically distributed)

$X_1, X_2 \overset{iid}{\sim}$ Bernoulli (1/3)

$g(X1, X2)$ function of two r.v.

a new r.v.
future insight 'total r.v.' aka 'sum r.v.'

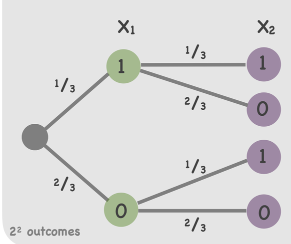
T_2

$T_2 = X_1 + X_2$ $Supp[T_2] = \{0,1,2\}$ (zero 1s, one 1s, two 1s)

$T_2 \sim \begin{cases} 0 & \text{w.p. } 4/9 \\ 1 & \text{w.p. } 2/9 + 2/9 = 4/9 \\ 2 & \text{w.p. } 1/9 \end{cases}$

disjoint events can be added
two sets are called 'disjoint' if their intersection is the empty set ∅

11
10
01
00



$P(A\cap B) = P(A)\cdot P(B)$

		T
$P(X_1=1, X_2=1) = 1/3 \cdot 1/3 = 1/9$		2
$P(X_1=1, X_2=0) = 1/3 \cdot 2/3 = 2/9$	4/9	1
$P(X_1=0, X_2=1) = 2/3 \cdot 1/3 = 2/9$		
$P(X_1=0, X_2=0) = 2/3 \cdot 2/3 = 4/9$		0

2^2 outcomes

T_3

$X_1,X_2,X_3 \overset{iid}{\sim}$ Bernoulli (1/3)

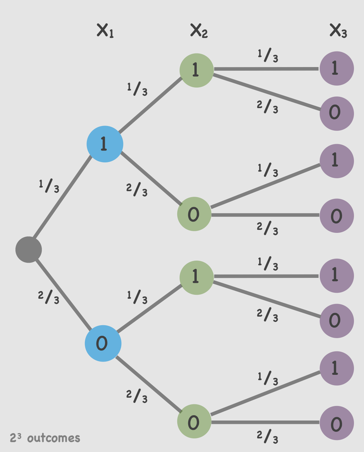
$T_3 = X_1 + X_2 + X_3$ $Supp[T_3] = \{0,1,2,3\}$ (zero 1s, one 1s, two 1s, three 1s)

$T_3 \sim \begin{cases} 0 & \text{w.p. } 8/27 \\ 1 & \text{w.p. } 4/27+4/27+4/27 = 12/27 \\ 2 & \text{w.p. } 2/27+2/27+2/27 = 6/27 \\ 3 & \text{w.p. } 1/27 \end{cases}$

$1\cdot(1/3)^0(2/3)^3$
$3\cdot(1/3)^1(2/3)^2$
$3\cdot(1/3)^2(2/3)^1$
$1\cdot(1/3)^3(2/3)^0$

disjoint events can be added

111
110
101
100
011
010
001
000



$P(A\cap B\cap C) = P(A)\cdot P(B)\cdot P(C)$

			T
$P(X_1=1, X_2=1, X_3=1) = 1/3\cdot1/3\cdot1/3 = 1/27$	or	$(1/3)^3 (2/3)^0$	3
$P(X_1=1, X_2=1, X_3=0) = 1/3\cdot1/3\cdot2/3 = 2/27$	or	$(1/3)^2 (2/3)^1$	2
$P(X_1=1, X_2=0, X_3=1) = 1/3\cdot2/3\cdot1/3 = 2/27$	or	$(1/3)^2 (2/3)^1$	2
$P(X_1=1, X_2=0, X_3=0) = 1/3\cdot2/3\cdot3/3 = 4/27$	or	$(1/3)^1 (2/3)^2$	1
$P(X_1=0, X_2=1, X_3=1) = 2/3\cdot1/3\cdot1/3 = 2/27$	or	$(1/3)^2 (2/3)^1$	2
$P(X_1=0, X_2=1, X_3=0) = 2/3\cdot1/3\cdot2/3 = 4/27$	or	$(1/3)^1 (2/3)^2$	1
$P(X_1=0, X_2=0, X_3=1) = 2/3\cdot2/3\cdot1/3 = 4/27$	or	$(1/3)^1 (2/3)^2$	1
$P(X_1=0, X_2=0, X_3=0) = 2/3\cdot2/3\cdot2/3 = 8/27$	or	$(1/3)^0 (2/3)^3$	0

2^3 outcomes

remember Pascal's Triangle ?

$\binom{3}{0}$ 000
$\binom{3}{1}$ 100 010 001
$\binom{3}{2}$ 110 011 101
$\binom{3}{3}$ 111

$1\cdot(1/3)^0(2/3)^3$
$3\cdot(1/3)^1(2/3)^2$
$3\cdot(1/3)^2(2/3)^1$
$1\cdot(1/3)^3(2/3)^0$

$T_3 \sim \binom{n}{x}(1/3)^x(2/3)^{3-x}$

PMF for Binomial (3, 1/3) — sample size, probability of success

3 n

1000 coin flips – 600 Heads

$P(600H) = P(X=600) = \binom{1000}{600}(1/2)^{600}(1/2)^{400}$

$X \sim$ Binomial (1000, 1/2) this is how it all connects

two ways to look at Binomial — both are the same mathematically

a bag so big that, if you take one ball out, the change in the ratio of success is negligible and observed as fixed OR a small bag but you are sampling with replacement thus keeping the ratio of success fixed

a really big bag with a fixed proportion of successes

$T = \lim_{N\to\infty} X \sim \text{Hypergeometric}(n,p,N)$

$T = X_1, X_2, \dots, X_n$ where $X_1, X_2, \dots, X_n \overset{iid}{\sim}$ Bernoulli (p)

you're summing a whole bunch of iid Bernoullis and at the end you ask – how many successes did I get ?

Binomial(n,p) == $X_1, X_2, \dots, X_n \overset{iid}{\sim}$ Bernoulli(p)

T_n, p

$X_1,X_2,\dots,X_n \overset{iid}{\sim}$ Bernoulli (p)

$T_n = X_1 + X_2 + \dots + X_n$ $Supp[T_n] = \{0,1,2,\dots,n\}$ (zero 1s, one 1s, two 1s, n 1s)

$T_n = \sum_{i=1}^{n} X_i$

1…11
1…10
1…01
1…00
…
0…11
0…10
0…01
0…00

$T_n \sim \begin{cases} 0 & \text{w.p. } \binom{n}{0} p^0 (1-p)^n \\ 1 & \text{w.p. } \binom{n}{1} p^1 (1-p)^{n-1} \\ 2 & \text{w.p. } \binom{n}{2} p^2 (1-p)^{n-2} \\ 3 & \text{w.p. } \binom{n}{3} p^3 (1-p)^{n-3} \\ \dots \\ n-2 & \text{w.p. } \binom{n}{n-2} p^{n-2} (1-p)^2 \\ n-1 & \text{w.p. } \binom{n}{n-1} p^{n-1} (1-p)^1 \\ n & \text{w.p. } \binom{n}{n} p^n (1-p)^0 \end{cases}$

2^n outcomes

PMF for ~ Binomial (n,p)

p 1/3 1–p 2/3

T_n

$X_1,X_2,\dots,X_n \overset{iid}{\sim}$ Bernoulli (1/3)

$T_n = X_1 + X_2 + \dots + X_n$ $Supp[T_n] = \{0,1,2,\dots,n\}$ (zero 1s, one 1s, two 1s, n 1s)

$T_n = \sum_{i=1}^{n} X_i$

1…11
1…10
1…01
1…00
…
0…11
0…10
0…01
0…00

$T_n \sim \begin{cases} 0 & \text{w.p. } \binom{n}{0} (1/3)^0 (2/3)^n \\ 1 & \text{w.p. } \binom{n}{1} (1/3)^1 (2/3)^{n-1} \\ 2 & \text{w.p. } \binom{n}{2} (1/3)^2 (2/3)^{n-2} \\ 3 & \text{w.p. } \binom{n}{3} (1/3)^3 (2/3)^{n-3} \\ \dots \\ n-2 & \text{w.p. } \binom{n}{n-2} (1/3)^{n-2} (2/3)^2 \\ n-1 & \text{w.p. } \binom{n}{n-1} (1/3)^{n-1} (2/3)^1 \\ n & \text{w.p. } \binom{n}{n} (1/3)^n (2/3)^0 \end{cases}$

2^n outcomes

PMF for ~ Binomial (n, 1/3)

CDF of Binomial

what is a CDF of a Binomial ?

$F(x) = P(X\le x)$

$F(x) = \sum_{i=0}^{x} p(x)$

$F(x) := P(X\le x) = \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i}$

incomplete regularized gamma function $= I_{1-p}(n-k, 1+k) := (n-k)\binom{n}{k}\int_0^{1-p} t^{n-k-1}(1-t)\,dt$

no closed form ! if you find one you get a Fields Medal

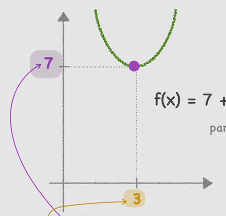
Geometric r.v. model

we keep trying and as soon as we succeed we stop

$X_1, X_2 \dots \overset{iid}{\sim}$ Bernoulli (p)

infinite series of binary experiments all with probability p of success where all are independent of one another (iid)
infinite sequence of iid r.v.'s

self-note:
think of the y-coordinates as indices and x-coordinates as 0s and 1s.
0==failure
1==success



$f(x) = 7 + (x-3)^2$ parabola shape

$\min(f(x)) = 7$
$\arg\min(f(x)) = 3$
$\max(f(x)) = $ dne
$\arg\max(f(x)) = $ dne

C++ actual parameter aka argument

keep on trying until $X_t=1$(success)

min() assigns to T the index-t at which 1(success) occurs, it could be 1 or n

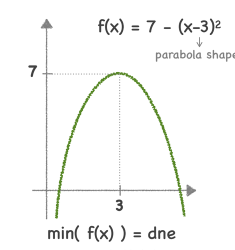
let $T = \min\{t: X_t=1\}$ if(X_t==1) return t

1 2 3 4 5 6 7
0 0 1 1 0 0 1 …
1st success

{ 3 , 4 , 7 , 11 , 18 , … }

first success 'stopping time'

$f(x) = 7 - (x-3)^2$ parabola shape



$\min(f(x)) = $ dne
$\arg\min(f(x)) = $ dne
$\max(f(x)) = 7$
$\arg\max(f(x)) = 3$

probability of stopping time at location 1 – meaning the success occurs at the first try

X: 1 (t: 1) done – experiment is over
$P(1) = P(T=1) = P(X_1=1) = p$ remember? p = Probability of Success

X: 0 1 (t: 1 2) stop
$P(2) = P(T=2) = P(X_1=0, X_2=1) = P(X_1)\cdot P(X_2) = (1-p)\cdot p$ — 1st success; independence of Bernoulli experiments; probability of failure · probability of success

X: 0 0 1 (t: 1 2 3) stop
$P(3) = P(T=3) = P(X_1=0, X_2=0, X_3=1) = P(X_1)\cdot P(X_2)\cdot P(X_3) = (1-p)\cdot(1-p)\cdot p = (1-p)^2\cdot p$

X: 0 0 0 … 1 (t: 1 2 3 … t; t–1) stop
$P(t) = P(T=t) = P(X_1=0, X_2=0, \dots, X_t=1) = P(X_1)\cdot P(X_2)\cdot\dots\cdot P(X_t) = (1-p)\cdot(1-p)\cdot\dots\cdot(1-p)\,p = (1-p)^{t-1}\cdot p$

$P(x) = P(T=x) = (1-p)^{x-1} p$ — x = # of trials

switching back to the standard r.v. notation T→X

$X \sim \text{Geometric}(p) := (1-p)^{x-1} p$ — p(x) – PMF

PMF

Parameter Space $p \in (0,1)$ probability of success lies between 0 and 1 (exclusively) b/c it comes from the underlying Bernoullis
if p=1 X~Deg(1) – you always win
if p=0 X~Deg(0) – you always loose

Supp[X] ? possible outcomes ? **Supp[X]** $= \{\cancel{0}, 1, 2, \dots\} = \mathcal{N}$
b/c you can't do something 0 times , you need at least 1

is this a legitimate r.v. ? $\sum_{x\in Supp[X]} p(x) = 1$

Geometric Supp[X] = { 1, … } → $\sum_{x=1}^{\infty}(1-p)^{x-1}p \overset{?}{=} 1$ | ÷p

$q := 1-p$
$p+q = 1$
$p = 1-q$
$p\in(0,1) \to q\in(0,1)$

$\sum_{x=1}^{\infty}(1-p)^{x-1} = 1/p$ | plug in for q

$\sum_{x=1}^{\infty} q^{x-1} = 1/(1-p)$ | **reindexing** x=1 to x=0

$\sum_{x=0}^{\infty} q^{0} = 1/(1-p)$

$S := \sum_{x=0}^{\infty} q^0 = \frac{1}{1-q}$ | express S in 'open form'

Geometric Series this is how this r.v. gets its name $\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}$

$S = q^0 + q^1 + q^2 + q^3 + \dots$ | anything^0 = 1
$= 1 + q^1 + q^2 + q^3 + \dots$
$= 1 + q\,\underbrace{(1 + q + q^2 + \dots)}_{S}$
$= 1 + qS$

$S - (qS) = 1$
$S(1-q) = 1$ $S = \frac{1}{1-q}$ QED

does CDF have a closed form ?

CDF is defined as $F(x) := P(X\le x) = \sum_{i=0}^{x}(1-p)^{i-1}p$ (CDF mathematically) Hard !!

OR

$F(x) := 1 - P(X>x)$

Complement Rule
$P(A) = 1-P(A^c)$
$P(A^c) = 1-P(A)$

$P(X>x) = P($ X: 0 0 1 … 0 | 0 1 … ; x: 1 2 3 … x x+1 $)$ — these are all failures: (1-p); success is somewhere here

$P(X>x) = (1-p)^x$

$F(x) := 1-(1-p)^x = P(X\le x)$
$P(X\le x) = 1 - P(X>x)$

CDF

Def: CDF F(x) $P(X\le x)$

seen another way – proof

$P(X>x) = P(X=x+1) + P(X=x+2) + P(X=x+3) + \dots$
$= \sum_{i=x+1}^{\infty}(1-p)^{i-1}p$ reindexing trick
$= \sum_{i=1}^{\infty}(1-p)^{x+i-1}p$
$= (1-p)^x \underbrace{\sum_{i=1}^{\infty}(1-p)^{i-1}p}_{=1}$

$P(X>x) = (1-p)^x$ a very simple closed form

probability of success occurring at x_{th} try OR sooner

probability of success occurring after the x_{th} try

You can think of the Geometric model as a sequence of iid Bernoulli(p) which could go on forever and you wait for the first one that hits success which is also the one that motivates your PMF.

you can think of this as the stopping time of $X_1, X_2, \dots \overset{iid}{\sim}$ Bernoulli(p)

e.g.

You play poker **until** you get a Royal Flush on which occasion you stop. On which hand do you get it for the 1st time?

$P(RF) = \frac{4}{\binom{52}{5}} = 1.53/\text{million} = .00000153$

$P(RF) = \frac{|RF|}{\binom{52}{5}} = \frac{\binom{4}{1}\binom{1}{1}}{\binom{52}{5}}$

- Build a r.v. **model** for a hand number in which you get that Royal Flush.
  this is usually a one liner

  $X_1, X_2, \dots \overset{iid}{\sim}$ Bernoulli(.00000153) stopping time

  $X \sim$ Geometric(.00000153) done

  X ~ Geometric(p)

- What is the probability you get a Royal Flush **on** the 1,000,000 hand ?

  Def: PMF p(x) $P(X=x)$

  $P(X=x) = (1-p)^{x-1}p$

  index=1,000,000 or on the 1,000,000 try; 1 – .00000153 ≈ .9999985

  $P(X=1{,}000{,}000) = (.9999985)^{999999}\cdot .00000153$

- What is the probability you get a Royal Flush **on** the 1,000,000 hand **or sooner**?

  CDF or later ? 1–CDF

  Def: CDF F(x) $P(X\le x)$

  $F(x) = P(X\le x) = 1-(1-p)^x$

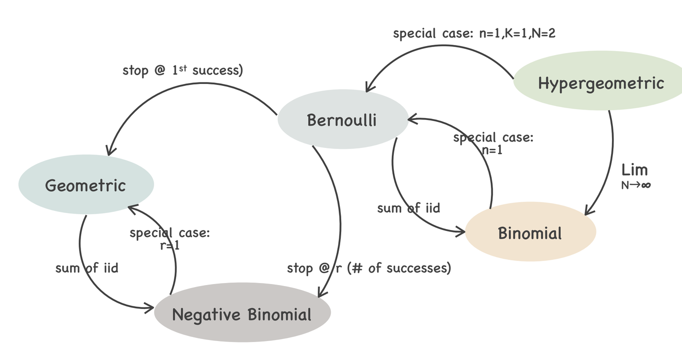
  $P(X\le 1{,}000{,}000) = F(1{,}000{,}000)$
  $= 1 - (.9999985)^{1{,}000{,}000}$
  $= .777 \approx 77\%$

  if you play poker million times there is 77% chance you'll get Royal Flush

- What is the probability you get a Royal Flush **after** the 1,000,000 hand ?

  $P(X>x) = (1-p)^x$

  $P(X>1{,}000{,}000) = (.9999985)^{1{,}000{,}000}$
  $= .223 \approx 22\%$



Philosophical Jump

$X \sim$ Bernoulli(p)

X=0 , X=1 → x=1 done

Supp[X]={0,1} but which one will it be? x=1 means it's that value. Little "x" means it's been "realized".

X refers to the model X (upper-case) represents the unrealized probability of X assuming a value in its support

x refers to the value that comes out of the model once the experiment is conducted and the numerical value is known X (upper-case) becomes x (lower-case) as soon as the experiment is over. Thus every experiment done on X spits out x equal to one of the values in the support of X

Realization (bring something abstract into reality)
Once the theoretical is realized it becomes material and is now set in stone. Realization converts the world of possibility into no-longer possible, thus eliminating the 'what if' into 'done'.
– must live in the Supp[X]

random variable r.v. : a model, not real (waiting to be happened)

datum : realization of a r.v. – once the coin is flipped we can see its outcome

data : realizations of r.v.'s

iid data : realizations of iid r.v.'s

upper-case X —conducting the experiment→ done lower-case x=1 realization

- 1 cup : 8 coins , 4 are marked , draw 3 out
  shake the cup, quarry 3 coins, how many are marked?

  $X \sim$ Hypergeometric(n,K,N) — 3,4,8 (sample size, # of success (marked), total population)

  every time we shake the cup and quarry 3 coins and see how many of them are marked – we make a realization

- 7 cups (in each) : 8 coins , 4 are marked , draw 3 out
  shake all 7 cups, quarry 3 coins, how many are marked?

  $X_1, X_2, X_3, X_4, X_5, X_6, X_7 \overset{iid}{\sim}$ Hypergeometric(n,K,N) — 3,4,8 (sample size, # of success (marked), total population)

  $x_1=1, x_2=2, x_3=2, x_4=3, x_5=1, x_6=2, x_7=2$

  Supp[X] = {0,1,2,3}

- 7 cups (in each) : 8 coins , draw 8 out
  shake all 7 cups, capsize all 7 cups, how many HEADS in each group ?

  $X_1, X_2, X_3, X_4, X_5, X_6, X_7 \overset{iid}{\sim}$ Binomial(n,p) — 8, 1/2 (sample size, probability of a head)

  $x_1=3, x_2=3, x_3=4, x_4=4, x_5=5, x_6=6, x_7=5$

  Supp[X] = {0,1,2,3,…,8}

- 1 cup : 8 coins , draw 8 out
  shake the cup, quarry 8 coins, how many HEADS?

  $X \sim$ Binomial(n,p) == $X_1\ X_2 \dots X_8 \overset{iid}{\sim}$ Bernoulli(1/2) — 8, 1/2 (sample size, probability of a head)

  Supp[X] = {0,1,2,3,…,8}