10

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

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

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

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

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

Binomial distribution Proof

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

verify $\sum_{x\in[0,...,n]} \binom{n}{x} p^x(1-p)^{n-x} = \boxed{1}$ if the Binomial it truly a r.v. we must prove = 1 how ?

Binomial Supp[X] = { 0,1, ... , 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
$$(p+(1-p))^n = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$
$$(p-p+1)^n =$$
$$(1)^n =$$
$$\boxed{1} =$$

QED

idd

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 independency ? $P(A|B) = P(A)$, $P(B|A) = P(B)$

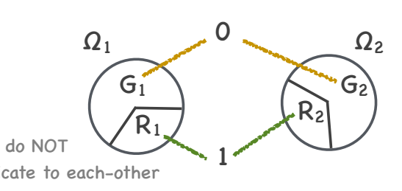
Multiplication Rule $P(B\cap A) = P(B)\cdot P(A)$

X_1 & X_2 are **independent & identically distributed - iid**

if X_1 and X_2 are independent & $X_1 \overset{d}{=} X_2$ denoted $X_1\ X_2 \overset{iid}{\sim}$

equal in distribution

(statistics) patients in a medical trial are considered iid



spinners do NOT communicate to each-other

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

$X_1 \sim$ 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.

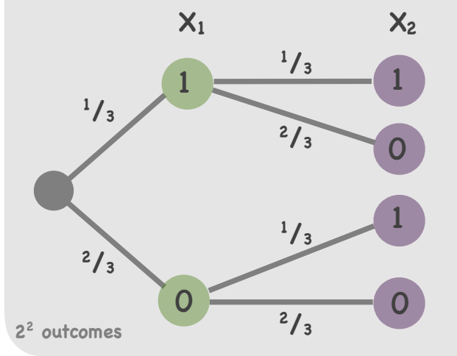
a new r.v.

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



11
10
01
00

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

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

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

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

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

T_3

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

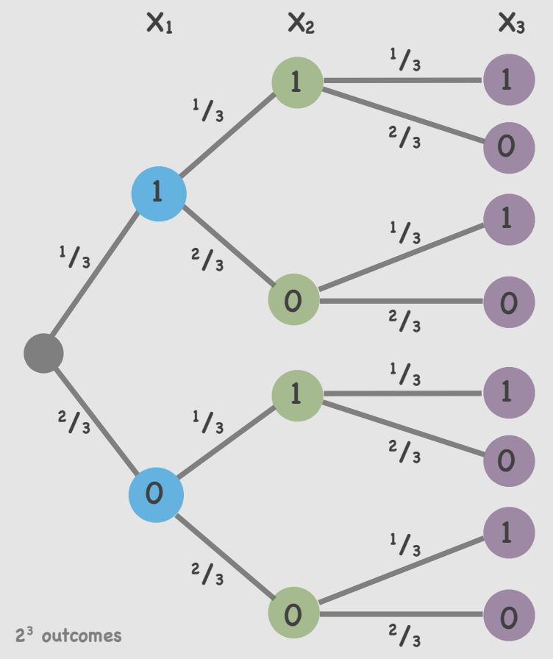
$T_3 = X_1 + X_2 + X_3$ $Supp[T_3] = \{0,1,2,3\}$ (zero 1s, one 1s, two 1s, three 1s)

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

disjoint events can be added

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

$T_3 \sim \binom{n}{x}(1/3)^x(2/3)^{3-x}$ → PMF for Binomial (3,1/3) (sample size, probability of success)



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

remember Pascal's Triangle ?

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

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

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

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

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

- $\lim_{N\to\infty} X \sim$ Hypergeometric (n,p,N)
- $X_1\ X_2 \dots X_n \overset{iid}{\sim}$ Bernoulli (p)

you're doing 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)

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

11...1	0	w.p. $\binom{n}{0} p^0 (1-p)^n$
11...0	1	w.p. $\binom{n}{1} p^1 (1-p)^{n-1}$
10...1	2	w.p. $\binom{n}{2} p^2 (1-p)^{n-2}$
10...0	3	w.p. $\binom{n}{3} p^3 (1-p)^{n-3}$
...	...	
01...1	n-2	w.p. $\binom{n}{n-2} p^{n-2} (1-p)^2$
01...0	n-1	w.p. $\binom{n}{n-1} p^{n-1} (1-p)^1$
00...1	n	w.p. $\binom{n}{n} p^n (1-p)^0$
00...0		

2ⁿ 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)

$T_n = X_1 + X_2 + \dots + X_n$ $Supp[T_n] = \{0,1,2,\dots,n\}$

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

11...1	0	w.p. $\binom{n}{0}(1/3)^0(2/3)^n$
11...0	1	w.p. $\binom{n}{1}(1/3)^1(2/3)^{n-1}$
10...1	2	w.p. $\binom{n}{2}(1/3)^2(2/3)^{n-2}$
10...0	3	w.p. $\binom{n}{3}(1/3)^3(2/3)^{n-3}$
...	...	
01...1	n-2	w.p. $\binom{n}{n-2}(1/3)^{n-2}(2/3)^2$
01...0	n-1	w.p. $\binom{n}{n-1}(1/3)^{n-1}(2/3)^1$
00...1	n	w.p. $\binom{n}{n}(1/3)^n(2/3)^0$
00...0		

2ⁿ outcomes

PMF for ~ Binomial (n ,1/3)

CDF of Binomial

what is a CDF of a Binomial ?

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

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

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

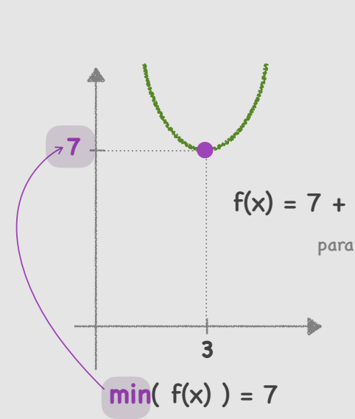
we keep trying and as soon as we succeed we stop

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

possibly infinite series of binary experiments with some probability independent of one another

infinite sequence of iid r.v.

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

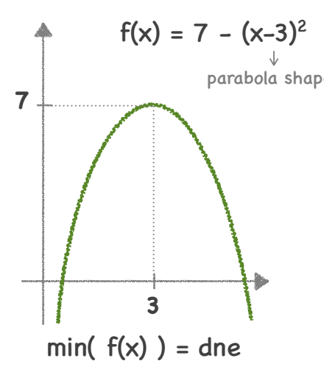
let $T = \min\{ t: X_t=1 \}$

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

$P(t) = P(T=t) = P(X_1=0, X_2=0, \dots, X_t=1) = P(X_1)\cdot P(X_2)\cdot\dots\cdot P(X_t) = (1-p)\cdot(1-p)\cdot\dots\cdot(1-p)\,p = (1-p)^{t-1}\cdot p$

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

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

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

Parameter Space ? $p\in(0,1)$ probability of success lies between 0 and 1 (exclusively) b/c it descends from Bernoulli
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

e.g.

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)

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?

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

- 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(p)

$X \sim$ Geometric(.00000153) done

- What is the probability you get a Royal Flush on the 1000,000 hand ?

index=1000,000 or on the 1000,000 try; $1 - .00000153 \approx .9999985$

$P(X=1000{,}000) = (.9999985)^{999999}\cdot .00000153$ — $P(X=x) = (1-p)^{x-1}p$

- What is the probability you get a Royal Flush on the 1000,000 hand or sooner? (CDF) or later ? (1-CDF)

$P(X\le 1000{,}000) = F(1000{,}000)$
$= 1 - (.9999985)^{1000{,}000}$
$= .777 \approx 77\%$

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

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

- What is the probability you get a Royal Flush after the 1000,000 hand ?

$P(X>1000{,}000) = (.9999985)^{1000{,}000}$
$= .223 \approx 22\%$

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

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 $\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}$

$S = q^0 + q^1 + q^2 + q^3 + \dots$ | anything⁰ = 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**

$P(X>x) = P(\ 0_1\ 0_2\ 1_3 \dots 0_x\ 0_{x+1}\ 1 \dots)$ these are all failures: (1-p); success is somewhere here

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

$F(x) := 1-(1-p)^x = P(X\le x)$ — probability of success occurring at x_{th} try OR sooner

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 after the x_{th} try

11

Negative Binomial r.v. model

we keep on trying and as soon as we succeed r-times we stop

Previously:
I'm going to wait till I hit the first success then I'll stop and return the number of steps (index at which the 1_{st} success occurred)

$X \sim \min\{ t: X_t = 1\}$

Now:
I'm going to wait till I hit r successes then I'll stop and return the number of steps (index at which the r_{th} success occurred)

$X \sim \min\{ t: \sum_{i=1}^{t} X_i = r\}$

X: 0 0 0 ... 1 ... 1 ... 0 0 1 ... (t: 1 2 3 ... ... t-2 t-1 t) r=3 r=3 stop, return t

Let's define a new r.v. which illustrates the above situation, assuming r=3 (we stop the experiment as soon as we get the 3rd success)

$P(X=0) =$ ✗ we can't get 3 successes in 0 tries
$P(X=1) =$ ✗ we can't get 3 successes in 1 trie
$P(X=2) =$ ✗ we can't get 3 successes in 2 tries

$P(X=3) = p\cdot p\cdot p = p^3$ we can get 3 successes in 3 tries
prob. of independent events are multiplied - think prob. tree

X: 1 1 1 (t: 1 2 3=t) r=3 r=3 stop, return t → X=t

$P(X=4) = (1-p)\cdot p\cdot p\cdot p$
$= p\cdot(1-p)\cdot p\cdot p$
$= p\cdot p\cdot(1-p)\cdot p$ } $= 3\cdot(1-p)\cdot p^3$
$= \underbrace{p\cdot p\cdot p}_{\text{done}}\cdot\cancel{(1-p)}$

$\binom{3}{1}$

$P(X=5) =$
0011 1
0101 1
0110 1
1001 1
1100 1
1010 1
} $= \boxed{6}\cdot(1-p)^2 p^3$ 5 tries = 2 failures and 3 successes

$\binom{4}{2}$

notation switch t → x

$P(X=x) = \binom{x-1}{2}(1-p)^{x-3}p^3$ we need 3 successes; failures = #of tries - 3 successes

the last try must be a success thus we are only concerned with x-1 preceding trials which contain the remaining two successes

X: 0 0 0 ... 1 ... 1 ... 0 0 1 ... (x: 1 2 3 ... ... x-2 x-1 x)

r-1 (total-last one); last r (success) — think of it as a single unit and not as r=# of successes thus = 1

x-1 (total-last one); last x — last index of success

# of failures: (x-1)-(r-1) = x-r

thus our new r.v. can be generalized as: drumroll!!

$X \sim \text{Negative Binomial}(r,p) := P(X=x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r$ (# of successes, prob. of success)

special case: if r=1 → Geometric PMF

$Supp[X] = \{r, r+1, \dots\}$ $r\in\mathcal{N}$

Parameter Space $p\in(0,1)$ - underlying Bernoulli

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

Neg Binomial Supp[X] = { r, ... }

$\sum_{x=r}^{\infty}\binom{x-1}{r-1}(1-p)^{x-r}p^r \overset{?}{=} 1$

Geometric Series $\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}$ if $a\in(0,1)$

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

$\sum_{x=1}^{\infty}(1-p)^{x-1} = p^{-1}$ | take 1st derivative both sides

$\sum_{x=2}^{\infty}(x-1)(1-p)^{x-2} = (-1)p^{-2}$ | take 2nd derivative both sides

$\sum_{x=3}^{\infty}(x-1)(x-2)(1-p)^{x-3} = (-1)(-2)p^{-3}$ | take 3rd derivative both sides

$\sum_{x=4}^{\infty}(x-1)(x-2)(x-3)(1-p)^{x-4} = (-1)(-2)(-3)p^{-4}$ | take **r-1** derivatives both sides

. . .

$\sum_{x=r}^{\infty}(x-1)(x-2)\dots(x-(r-1))(1-p)^{x-r} = \underbrace{(-1)(-2)(-3)\dots(r-1)}_{(r-1)!}p^{-r}$ (r-1)+1

$6\cdot5\cdot4\cdot3 = \frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{2\cdot1} = \frac{6!}{2!}$

$(x-1)\dots(x-(r-1)) = \frac{(x-1)(x-2)\dots(x-(r-1))(x-r)(x-(r+1)(x-(r+2))\dots\cdot2\cdot1}{(x-r)(x-(r+1)(x-(r+2))\dots\cdot2\cdot1} = \frac{(x-1)!}{(x-r)!}$

$\sum_{x=r}^{\infty}\frac{(x-1)!}{(x-r)!}(1-p)^{x-r} = (r-1)!\ p^{-r}$ | ÷ (r-1)!

$\sum_{x=r}^{\infty}\frac{(x-1)!}{(x-r)!\,(r-1)!}(1-p)^{x-r} = p^{-r}$ | (r-1)! = ((x-1)-(x-r))! = (x-1-x+r)!

translation:
(one less # of successes)! = (one less # of trials - all failures)!
r = # of successes
x = total # of trials

$\sum_{x=r}^{\infty}\frac{(x-1)!}{(x-r)!\,((x-1)-(x-r))!}(1-p)^{x-r} = p^{-r}$

$\frac{n!}{k!\,(n-k)!} = {}_nC_k = \binom{n}{k}$

$\sum_{x=r}^{\infty}\binom{x-1}{x-r}(1-p)^{x-r} = p^{-r}$ | · p^r

$\sum_{x=r}^{\infty}\binom{x-1}{x-r}(1-p)^{x-r}p^r = 1$ | $p^r/p^r = 1$

QED

Equivalent Parametrization

$X \sim$ Negative Binomial(r,p)
Let x = # of failures

X: 0 0 0 ... 1 ... 1 ... 0 0 1 ... (i: 1 2 3 ... ... i-2 i-1 i) index

r = total # of success; r-1 (total-last one); last r (success) — think of it as a single unit and not as r=total # of success, thus here just 1 success

x+r-1 (total-last one) — # of failures + (# of successes - 1)

$P(X=x) = \binom{x+r-1}{r-1}(1-p)^x p^r$ — out of (# of failures) + (# of successes - 1) choose (# of successes - 1) — $\frac{6!}{2!(6-2)!} = \frac{6\cdot5}{2} = 15$

=

$P(X=x) = \binom{x+r-1}{x}(1-p)^x p^r$ — out of (# of failures) + (# of successes - 1) choose (# of failures) — $\frac{6!}{4!(6-4)!} = \frac{6\cdot5}{2} = 15$

$\frac{(x+r-1)!}{(r-1)!\,[(x+\cancel{(r-1)})-\cancel{(r-1)}]!} \longrightarrow \frac{(x+r-1)!}{(r-1)!\,x!} = \frac{(x+r-1)(x+r-2)\cdot\dots\cdot r}{x!} = (-1)^x\frac{(-r)(-r-1)(-r-2)\cdot\dots\cdot(-r-x+1)}{x!}$
$= (-1)^x\binom{-r}{x}$

think of r-1 as a single unit

this is where Negative Binomial gets its name

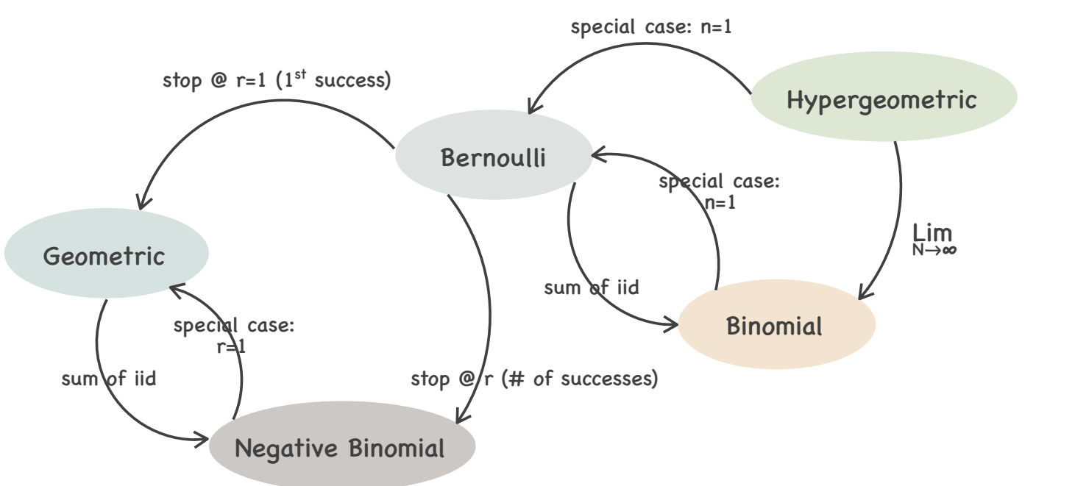
$p(x) = (-1)^x\binom{-r}{x}(1-p)^x p^r$

Roll a die until you get 17 6's. What is the prob. it takes until the 107th roll?

$X \sim$ Negative Binomial(17,1/6) (# of successes, prob. of successes)

$P(X=107) = \binom{106}{16}(5/6)^{90}(1/6)^{17}$

we'll prove this later =
- $X_1, X_2, \dots, X_r \overset{iid}{\sim}$ Geometric(p)
- $X_1 + X_2 + \dots + X_r \overset{iid}{\sim}$ Negative Binomial(r,p)



Philosophical Jump

$X \sim$ Bernoulli(p) X refers to the model (the model is technically a function) that wires abstract to numerical

X=0 , X=1 x something that comes out of the model once the experiment is conducted and the numerical value is known

Realization (making fantasy into reality)
Once the theoretical is realized it becomes material and in 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

$X \xrightarrow{\text{conducting the experiment}}$ done 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}

- 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}

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

$X_1, X_2, X_3, X_4, X_5, X_6, X_7 \overset{iid}{\sim}$ 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}