October 13, 2016

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

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

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

$X \sim \text{Hypergeometric}(n, p, N) := \dfrac{\binom{pN}{x}\binom{(1-p)N}{n-x}}{\binom{N}{n}}$ $\left\{ \begin{array}{l} p = \frac{K}{N} \\ K = pN \end{array} \right\}$ Equivalent Parameterization

$\lim_{N \to \infty} \text{Hyper}(n, p, N) = \text{Binom}(n, p)$

Recall the following fact:

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

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

Recall: $\underbrace{(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}}$ } Binomial Theorem

Let $a = p$, $b = 1-p$, $i = x$

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

$$1^n$$

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

This is why the binomial r.v. got its name - from binomial theorem.

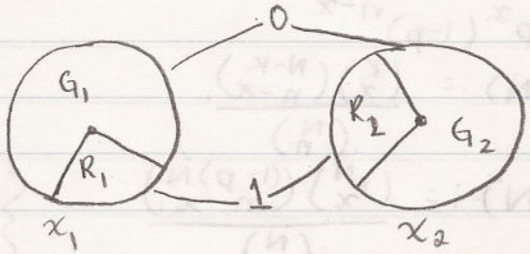
X_1 and X_2 are independent random variables if ...

Joint Mass Function (JMF)
$$\left.\begin{array}{l} P(X_1 = x_1 \mid X_2 = x_2) = P(X_1 = x_1) \text{ or} \\ P(X_2 = x_2 \mid X_1 = x_1) = P(X_2 = x_2) \text{ or} \\ P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)\, P(X_2 = x_2) \end{array}\right\} \forall x_1 \in Supp[X_1], \forall x_2 \in Supp[X_2]$$

iid

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

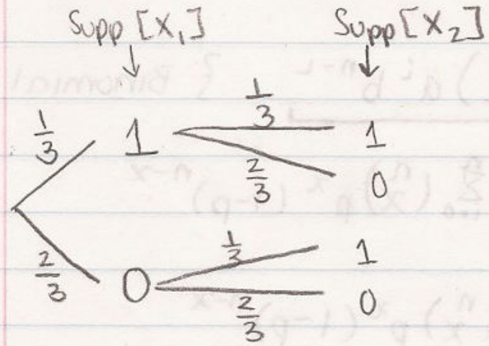
Example: $X_1, X_2 \overset{iid}{\sim} \text{Bern}(\frac{1}{3})$



A new r.v.
$T_2 = X_1 + X_2$
$= g(X_1, X_2)$
A function of two r.v.

→ How is T_2 distributed? $T_2 \sim$?

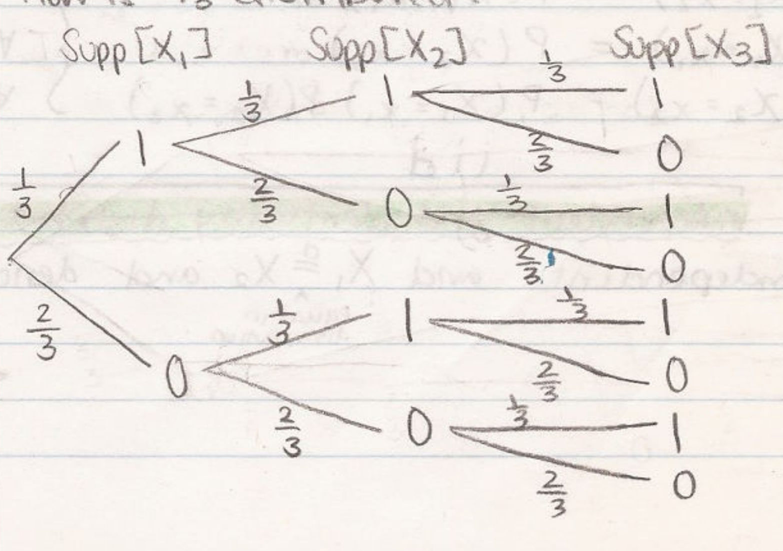
- What is the support of T? $\text{Supp}[T_2] = \{0, 1, 2\}$
  Because T_2 can be $0+1$, $0+0$, $1+0$, $1+1$.
- In order to find $T_2 \sim$, you have to find the prob of 0, 1, and 2. This is the PMF.



		$T_2 = X_1 + X_2$
$P(X_1 = 1, X_2 = 1)$	$= \frac{1}{9}$	2
$P(X_1 = 1, X_2 = 0)$	$= \frac{2}{9}$	1
$P(X_1 = 0, X_2 = 1)$	$= \frac{2}{9}$	1
$P(X_1 = 0, X_2 = 0)$	$= \frac{4}{9}$	0
	$\frac{9}{9}$	

→ $T_2 \sim \begin{cases} 0 & \text{wp } \frac{4}{9} \\ 1 & \text{wp } \frac{4}{9} \\ 2 & \text{wp } \frac{1}{9} \end{cases}$

→ How is T_3 distributed?



	$T_3 = X_1 + X_2 + X_3$
$(\frac{1}{3})^3 (\frac{2}{3})^0$	3
$(\frac{1}{3})^2 (\frac{2}{3})^1$	2
$(\frac{1}{3})^2 (\frac{2}{3})^1$	2
$(\frac{1}{3})^1 (\frac{2}{3})^2$	1
$(\frac{1}{3})^2 (\frac{2}{3})^1$	2
$(\frac{1}{3})^1 (\frac{2}{3})^2$	1
$(\frac{1}{3})^1 (\frac{2}{3})^2$	1
$(\frac{1}{3})^0 (\frac{2}{3})^3$	0

3rd Row of Pascal's Δ!

$$\rightsquigarrow T_3 \sim \begin{cases} 0 & \text{wp } 1\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^3 \\ 1 & \text{wp } 3\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^2 \\ 2 & \text{wp } 3\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^1 \\ 3 & \text{wp } 1\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^0 \end{cases} \qquad \begin{cases} 1 = \binom{3}{0} \\ 3 = \binom{3}{1} \\ 3 = \binom{3}{2} \\ 1 = \binom{3}{3} \end{cases}$$

!!! Pattern: $T_3 \sim \binom{3}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{3-x} = \text{Binom}\left(3, \frac{1}{3}\right)$!!!

$1\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^3$	$3\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^2$	$3\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^1$	$1\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^0$
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
0 0 0	1 0 0	1 1 0	1 1 1
	0 1 0	1 0 1	
	0 0 1	0 1 1	
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
$\binom{3}{0}$	$\binom{3}{1}$	$\binom{3}{2}$	$\binom{3}{3}$

$\rightsquigarrow X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Bern}\left(\frac{1}{3}\right)$

$\text{Supp}[T_n] = \{0, 1, \ldots, n\}$

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

$$T_n \sim \text{Binom}\left(n, \tfrac{1}{3}\right) \sim \begin{cases} 0 & \text{wp } \binom{n}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^n \\ 1 & \text{wp } \binom{n}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^{n-1} \\ 2 & \text{wp } \binom{n}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^{n-2} \\ \vdots & \vdots \\ n-2 & \text{wp } \binom{n}{n-2}\left(\frac{1}{3}\right)^{n-2}\left(\frac{2}{3}\right)^2 \\ n-1 & \text{wp } \binom{n}{n-1}\left(\frac{1}{3}\right)^{n-1}\left(\frac{2}{3}\right)^1 \\ n & \text{wp } \binom{n}{n}\left(\frac{1}{3}\right)^n\left(\frac{2}{3}\right)^0 \end{cases}$$

$\rightsquigarrow X_1, X_2, \ldots X_n \overset{iid}{\sim} \text{Bern}(p)$

$$T_n \sim \text{Binom}(n, p) \sim \begin{cases} 0 & \text{wp } \binom{n}{0}(p)^0(1-p)^n \\ 1 & \text{wp } \binom{n}{1}(p)^1(1-p)^{n-1} \\ 2 & \text{wp } \binom{n}{2}(p)^2(1-p)^{n-2} \\ \vdots & \\ n-2 & \text{wp } \binom{n}{n-2}(p)^{n-2}(1-p)^2 \\ n-1 & \text{wp } \binom{n}{n-1}(p)^{n-1}(1-p)^1 \\ n & \text{wp } \binom{n}{n}(p)^n(1-p)^0 \end{cases}$$

2 Concepts for Binomial:

① $\lim_{N\to\infty}$ Hyper (n, p, N) = Binom (n, p)

OR

② $X_1, X_2, \ldots, X_n \overset{iid}{\sim}$ Bern (p) = $X_1 + \ldots + X_n \sim$ Binom(n, p)

- Sampling with replacement and sampling without replacement are the same thing as long as the "bag" is big.

1000 Coin Flips

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

↳ $X \sim$ Binom $(1000, \frac{1}{2})$

CDF of X

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

What is the CDF of Binomial?

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

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

↳ Incomplete regularized gamma function

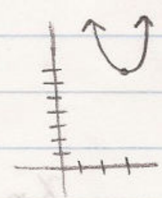
Consider the following: $X_1, X_2, \ldots \overset{iid}{\sim}$ Bern(p)

This is an infinite sequence of random variables.

Infinite experiments.

Calculus Stuff

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



- $\min(f(x)) = 7$
  - ↳ The lowest the function could get on the y axis.
- $\text{argmin}(f(x)) = 3$
  - ↳ Argument to a function. What can you put into f to get the $\min(f(x))$?

Back to $X_1, X_2, \ldots \overset{iid}{\sim} \text{Bern}(p)$

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

↳ "first success", "stopping time", The first time a success occurs.

$P(1) \Rightarrow P(T=1) \Rightarrow P(X_1=1) = p$

↳ 1 _ _ _ _ _

$P(2) \Rightarrow P(T=2) \Rightarrow P(X_1=0, X_2=1) \Rightarrow P(X_1=0) * P(X_2=1) = (1-p)p$

↳ 0 1 _ _ _

$P(3) \Rightarrow P(T=3) \Rightarrow P(X_1=0, X_2=0, X_3=1) = (1-p)^2 p$

↳ 0 0 1 _ _

⋮

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

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

Parameter Space

$p \in (0,1)$ ← Same parameter space as Bern.

Support

$\text{Supp}[X] = \{1, 2, 3, \ldots\} = \mathbb{N}$

↳ What are the possible outcomes? How many experiments would it take to get a succes? Can't be 0 b/c if you do 0 experiments, there's no way of getting a success.

True or False? Is $\sum_{x \in \text{Supp}[X]} p(x) = 1$ equal to 1? $\sum_{x=1}^{\infty} (1-p)^{x-1} p \overset{?}{=} 1$ True, it is a random variable

Proof:

Divide both sides by p. $\sum_{x=1}^{\infty} (1-p)^{x-1} = \frac{1}{p}$

Sub: $q = 1-p$ $\sum_{x=1}^{\infty} q^{x-1} = \frac{1}{1-q}$

Reindexing Trick $\sum_{x=0}^{\infty} q^{x} = \frac{1}{1-q}$

Sub: $\sum_{x=0}^{\infty} q^x = S$

$S = 1 + q^1 + q^2 + q^3 + \ldots$

$S = 1 + q(1 + q + q^2 + \ldots)$

$S = 1 + q(S)$

$S - q(S) = 1$

$S(1-q) = 1$

$S = \frac{1}{1-q}$

$\sum_{x=0}^{\infty} q^x = \frac{1}{1-q}$

$\left.\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}\right\}$ Geometric Series

↳ $q \in (0,1)$

↳ How $X \sim \text{Geom}$ got its name

CDF

$F(x) = P(X \le x) = \sum_{i=1}^{x} (1-p)^{i-1} p$ → closed form.

OR

$F(x) = 1 - P(X > x)$ → Complement Rule

$P(X > x) = P(\underset{1}{0}\ \underset{2}{0}\ \underset{3}{0}\ \underset{4}{0}\ \underset{5}{0} \ldots \underset{x}{0}\ \underset{x+1}{0}\ \underset{x+2}{1}\ \underset{x+3}{0} \ldots)$

All failures = 0. Success lies somewhere here.

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

so $F(x) = 1 - (1-p)^x$

Proof:

$P(X > x) = P(X = x+1) + P(X = x+2) + P(X = x+3) + \ldots$

$= \sum_{i=x+1}^{\infty} (1-p)^{i-1} p$

$= \sum_{i=1}^{\infty} \underbrace{(1-p)^{x+i-1}}_{} p$

looks like $(1-p)^x (1-p)^{i-1}$.

Pull out the constant(s).

$= (1-p)^x \underbrace{\sum_{i=1}^{\infty} (1-p)^{i-1} p}_{= 1}$ (We proved this to be true already).