Math 340/640 Lec 2

Review of 241 concept of rv

X vs x ← realization or datum

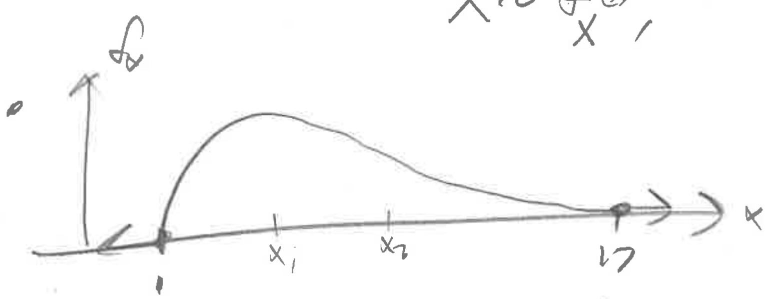
"data" = realizations from rv

- Coin flip

$X \sim f_X(x)$, continuous rv

One realization: one universe amongst a multiverse

$X_1, \ldots, X_n \overset{iid}{\sim} f_X(x)$



← S_X →

$[1, 17]$

$X \sim \text{Bern}(p)$, $S_X = \{0, 1\}$, $p \in (0, 1)$

named after Jacob Bernoulli (1600's)

$X \sim \text{Deg}(c)$, $S_X = \{c\}$, $c \in \mathbb{R}$

$X \sim \text{Unif}(\{a_1, a_2\}) = \begin{cases} a_1 & \text{wp } \frac{1}{2} \\ a_2 & \text{wp } \frac{1}{2} \end{cases}$

$S_X = \{a_1, a_2\}$, $a_1, a_2 \in \mathbb{R}$, $a_1 \neq a_2$

$X \sim \text{Unif}(A)$, $S_X = A$, $A \subset \mathbb{R}$

$= \frac{1}{|A|} \mathbb{1}_{x \in A} = p(x)$ "uniform discrete rv"

Convolutions $\quad T = X_1 + X_2$

$$p(t) = \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \mathbb{R}} p_{X_1, X_2}(x_1, x_2) \mathbb{1}_{t = x_1 + x_2}$$

$x_2 = t - x_1$

$x_2 \in \{t - x_1\}$ singleton set!

$$= \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \{t - x_1\}} p_{X_1, X_2}(x_1, x_2)$$

$$= \sum_{x_1 \in \mathbb{R}} p_{X_1, X_2}(t - x_1, x_2)$$

$$= \sum_{x \in \mathbb{R}} p_{X_1, X_2}(t - x, x)$$

$$p(T=t) = \sum_{x \in \mathbb{R}} p_{X_1,X_2}(x, t-x) \quad \text{general conv. formula}$$

if $X_1, X_2 \overset{ind}{\sim}$

$$= \sum_{x \in \mathbb{R}} p_{X_1}(x)\, p_{X_2}(t-x) = \sum_{x \in S_{X_1}} p^{old}_{X_1}(x)\, p_{X_2}(t-x)\, \mathbb{1}_{t-x \in S_{X_2}}$$

if $X_1, X_2 \overset{iid}{\sim}$

$$= \sum_{x \in \mathbb{R}} p(x)\, p(t-x) = \underbrace{\sum_{x \in S_X} p^{old}(x)\, p^{old}(t-x)\, \mathbb{1}_{t-x \in S_X}}_{\text{famous conv. formula}}$$

Let's use the famous conv. formula to solve for the PMF of $T = X_1 + X_2$ where $X_1, X_2 \overset{iid}{\sim} \text{Bern}(p)$

$$T \sim \sum_{x \in \{0,1\}} \left(p^x (1-p)^{1-x}\right)\left(p^{t-x}(1-p)^{\overbrace{1-(t-x)}^{1-t+x}}\right) \mathbb{1}_{t-x \in \{0,1\}}$$

$$= \sum_{x \in \{0,1\}} p^t (1-p)^{2-t}\, \mathbb{1}_{t \in \{x, x+1\}} = p^t(1-p)^{2-t} \sum_{x \in \{0,1\}} \mathbb{1}_{t \in \{x,x+1\}}$$

$$= p^t (1-p)^{2-t} \left(\mathbb{1}_{t \in \{0,1\}} + \mathbb{1}_{t \in \{1,2\}}\right)$$

$\binom{n}{k} := \frac{n!}{k!(n-k)!}\mathbb{1}_{k \in \{0,1,\ldots,n\}}\mathbb{1}_{n \in \mathbb{N}_0}$

Recall... ↗

$$\mathbb{1}_{t \in \{0,1\}} + \mathbb{1}_{t \in \{1,2\}} = \begin{cases} 1 & \text{if } t=0 \\ 2 & \text{if } t=1 \\ 1 & \text{if } t=2 \\ 0 & \text{o/t} \end{cases} = \binom{2}{t} := \frac{2!}{t!(2-t)!}\mathbb{1}_{t \in \{0,1,2\}}$$

$$\Rightarrow T \sim \binom{2}{t} p^t (1-p)^{2-t} = \text{Binom}(2, p) \quad \text{"Binomial rv"}$$

Let's do this again using combinations

$X_1, X_2 \overset{iid}{\sim} Bern(p) = \binom{1}{x} p^x (1-p)^{1-x}$

$\binom{1}{x} = \frac{1!}{x!(1-x)!} \mathbb{1}_{x \in \{0,1\}}$

$= \begin{cases} 1 & \text{if } x = 1 \\ 1 & \text{if } x = 0 \\ 0 & \text{o/w} \end{cases} = \mathbb{1}_{x \in \{0,1\}}$

$T_2 \sim \sum_{x \in \mathbb{R}} p(x)\, p(t-x) = \sum_{x \in \mathbb{R}} \Big(\overbrace{\binom{1}{x}}^{\mathbb{1}_{x\in\{0,1\}}} p^x (1-p)^{1-x} \binom{1}{t-x} p^{t-x} (1-p)^{1-t+x} \Big)$

$= \sum_{x \in \{0,1\}} p^t (1-p)^{2-t} \binom{1}{t-x}$

$= p^t (1-p)^{2-t} \left(\binom{1}{t} + \binom{1}{t-1} \right)$

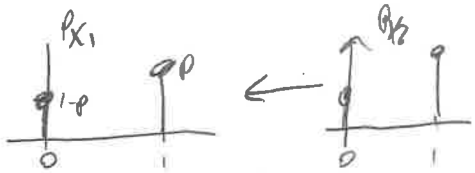
$= \binom{2}{t} p^t (1-p)^{2-t}$

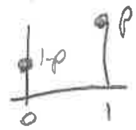
Pascal's Identity

$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}$

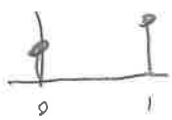
Why is it called a "convolution"?

"Convolve" means to roll or coil together, entwine.

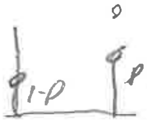




$\Rightarrow$ $t = 1+0 = 1$ w.p. $p(1-p)$



$\Rightarrow$ $t = 0+0 = 0$ w.p. $(1-p)^2$ AND $t = 1+1 = 2$ w.p. p^2



$\Rightarrow$ $t = 0+1 = 1$ w.p. $(1-p)\,p$

$$\Rightarrow p_T(t) = \begin{cases} 0 & \text{w.p. } (1-p)^2 \\ 1 & \text{w.p. } 2p(1-p) \\ 2 & \text{w.p. } p^2 \end{cases}$$

5

$X_1, X_2, X_3 \overset{iid}{\sim} Bern(p) \quad T_3 = X_1 + X_2 + X_3 \sim p(t) = ?$

Can we use the same formula? Yes, we know $T_2 = X_1 + X_2 \sim Binom(2, p)$

let $T_3 = X_3 + T_2 \sim \sum_{x \in \mathbb{R}} p_{X_3}(x)\, p_{T_2}(t-x)$

$$= \sum_{x \in \mathbb{R}} \binom{1}{x} p^x (1-p)^{1-x} \binom{2}{t-x} p^{t-x} (1-p)^{2-t+x}$$

$$= \sum_{x \in \{0,1\}} p^t (1-p)^{3-t} \binom{2}{t-x}$$

$$= p^t (1-p)^{3-t} \left(\binom{2}{t} + \binom{2}{t-1} \right)$$

Pascal's Δ identity

$$= \binom{3}{t} p^t (1-p)^{3-t} = Binom(3, p)$$

HW: use induction to prove general Binomial PMF.

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

$S_X = \{0, 1, \ldots, n\}$

$n \in \mathbb{N}, \ p \in (0,1)$

$X_1 \sim Binom(n, p)$ ind of $X_2 \sim Binom(m, p)$

$X_1 + X_2 \sim Binom(n+m, p)$. Need Vandermonde's combinatorial identity to prove but o/t simple. HW problem similar.

Why is it called the called Binomial?

Hyper Dingsy Identity $\quad \underbrace{\sum_{x \in \mathbb{R}}}_{x \in \{1, \ldots, n\}} \binom{n}{x} p^x (1-p)^{n-x} = p + (1-p) = 1$

by Binomial thm

possibly infinite seq. of rv's!

$B_1, B_2, B_3, \ldots \overset{iid}{\sim} \text{Bern}(p)$

"# failures before first success"

Let $X :=$ # of 0's before the first 1 occurs

$= \min_t \{t : B_t = 1\}$

Is X a rv?

$P(X=0) = P(B_1 = 1) = p$

$P(X=1) = P(B_1 = 0, B_2 = 1) = (1-p)p$

$P(X=2) = P(B_1=0, B_2=0, B_3=1) = (1-p)^2 p$

geometric r.v.

$\vdots$

$S_X = \mathbb{N}_0,\ p \in (0,1)$

$p(x) = P(X=x) = P(B_1=0, \ldots, B_x = 0, B_{x+1} = 1) = (1-p)^x p \, \mathbb{1}_{x \in \{0,1,\ldots\} = \mathbb{N}_0}$

$\text{Geom}(p) :=$

Why is it called the geometric rv?

Happy Dappy identity:

$$\sum_{x \in \mathbb{R}} p(x) = \sum_{x \in \mathbb{N}_0} (1-p)^x p$$

geometric series formula

$$= p\left(\frac{1}{1-(1-p)}\right) = \frac{p}{p} = 1 \checkmark$$

$$\sum_{i \in \mathbb{N}_0} q^i = \frac{1}{1-q} \quad \text{if } |q| < 1$$

let $X_1, X_2 \overset{iid}{\sim} \text{Geom}(p)$, let $T_2 = X_1 + X_2 \sim p(t) = ?$

failures before 2^{nd} success

$S_{T_2} = \{0, 1, \dots\}$

$$p(t) = \sum_{x \in S_X} \overset{iid}{p}(x)\, \overset{iid}{p}(t-x)\, \mathbb{1}_{t-x \in S_X}$$

want to use intersection identity

$$= \sum_{x \in \{0,1,\dots\}} (1-p)^x p\, (1-p)^{t-x} p\, \mathbb{1}_{t-x \in \{0,1,\dots\}}$$

$$t-x \in \{0,1,\dots\}$$
$$x-t \in \{0,-1,-2,\dots\}$$
$$x \in \{t, t-1, t-2, \dots\}$$

otherwise set intersection is zero!

$$= p^2 (1-p)^t \mathbb{1}_{t \in \mathbb{N}_0} \sum_{x \in \{0,1,\dots,t\}} 1 = (t+1)\, p^2 (1-p)^t\, \mathbb{1}_{t \in \mathbb{N}_0} = \text{NegBin}(2, p)$$

$$(t+1) = \binom{t+1}{1}$$

"Negative Binomial" r.v.
named after a definition that is irrelevant for this course

$X_1, X_2, X_3 \overset{iid}{\sim} \text{Geom}(p)$, let $T_3 = X_1 + X_2 + X_3 = X_3 + T_2 \sim p(t) = ?$

$$p(t) = \sum_{x \in S_X} (1-p)^x p\, (t-x+1)\, p^2 (1-p)^{t-x}\, \mathbb{1}_{t-x \in S_{T_2}}$$

$$= p^2 (1-p)^t \sum_{x \in \{0,1,\dots\}} (t-x+1)\, \mathbb{1}_{t-x \in \{0,1,\dots\}}$$

$$x \in \{t, t-1, t-2, \dots\}$$

$$= p^2 (1-p)^t \sum_{x \in \{0,1,\dots,t\}} (t-x+1) = \binom{t+2}{2} p^2 (1-p)^t = \text{NegBin}(3, p)$$

$$t \sum_{x \in \{0,1,\dots,t\}} (1) - \sum_{x \in \{0,1,\dots,t\}} x + \sum_{x \in \{0,1,\dots,t\}} (1)$$
$$= t(t+1) - \frac{t(t+1)}{2} + (t+1)$$

$$(t+1)^2 - \frac{t^2}{2} - \frac{t}{2}$$
$$= t^2 + 2t + 1 - \frac{t^2}{2} - \frac{t}{2}$$
$$= \frac{t^2 + 3t + 2}{2}$$
$$= \frac{(t+2)(t+1)}{2}$$
$$= \frac{(t+2)!}{2!\, t!} = \binom{t+2}{2}$$

MA students on HW will show that

r successes

$$X_1, X_2, \dots, X_r \overset{iid}{\sim} \text{Geom}(p) \implies T = X_1 + \dots + X_r \sim \text{NegBin}(r,p) = \binom{t+r-1}{r-1} p^r (1-p)^t$$

counting the zeroes when

Why does this PMF make sense? T is a wary for r successes. If the realization is t, then

$$\underbrace{\underset{1^{st}}{0}\ \underset{2^{nd}}{1}\ 0\ 1\ \dots\ 1\ 0}_{t+r-1}\ \underbrace{\underset{t+r^{th}}{1}}_{1}$$

r th success

r-1 successes

t zeroes

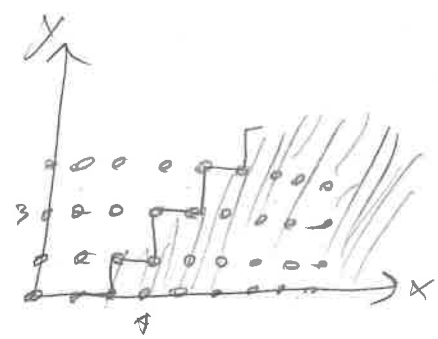
t zeroes

\# of ways to have r-1 successes in t+r-1 Bernoulli trials

SKIP Poisson rv for now

Another use of indicator functions

Let $X, Y \overset{iid}{\sim} \text{Geom}(p)$



$$P(X > Y) = \sum_{x \in S_X} \sum_{y \in S_Y} p_{X,Y}(x,y)\, \mathbb{1}_{x>y}$$

by iid

$$= \sum_{x \in \mathbb{N}_0} \sum_{y \in \mathbb{N}_0} p(x)\, p(y)\, \mathbb{1}_{x \geq y+1}$$

$$= \sum_{y \in \mathbb{N}_0} \sum_{x \in \mathbb{N}_0} (1-p)^x p\, (1-p)^y p\, \mathbb{1}_{x \geq y+1}$$