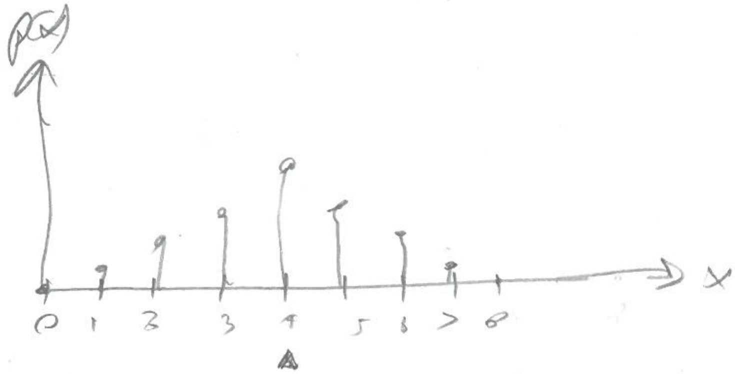
Lec 13 Math 241 10/24/17   Do n $X_1, \dots, X_8 \sim Bin(8, \frac{1}{2})$ calc. $\bar{x}$

$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ avg. r.v. $\bar{x} = \frac{x_1 + \dots + x_n}{n}$

realization fun

Where does $\bar{x} \to$? 4?

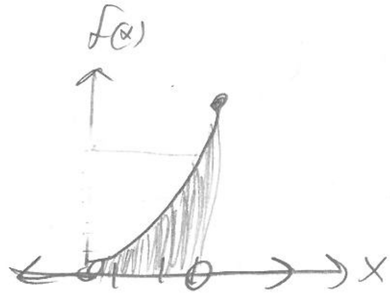


pivot pt
or
balance pt.

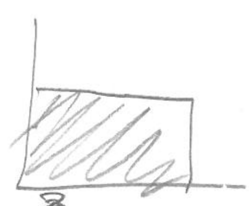
The # 4 "sums up" $X \sim Bin(n,p)$. Do you agree?

Recall calculus...

$f(x) = x^2$ where $x \in [0,3)$



$f(x) = 3, \; x \in [0,3)$



$\int_0^3 f(x)\,dx = 9$

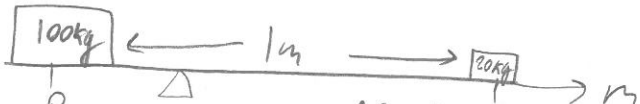
$\int_0^3 f(x)\,dx = 9$ What is an integral? A function of a function...

$G[f] = \int_0^3 f(x)\,dx$ $\qquad G:$ functions $\longrightarrow \mathbb{R}$ (a single #)

functional function "A function of a function".

It appears $\bar{X} \longrightarrow$ the "pivot pt" or the "balance pt"!

HS physics



Where is the balance pt? Need Angular Momentum $= 0$

$$\sum_i w_i(d_i - d^*) = 0 \Rightarrow \sum w_i d_i - \sum w_i d^* = 0 \Rightarrow \frac{\sum w_i d_i}{\sum w_i} = d^*$$

$$= \frac{0\cdot 100 + 1\cdot 20}{100+20} = 0.167\,m$$

Let's use same principle here... Weights = probs distance $= X$ $\quad \frac{\sum_i p_i x_i}{\sum_i p_i} = x^*$ $\quad p(x)$ $\quad 1$

interchangeable

$\mathbb{E}[X] := \mu := \sum_{x \in \text{supp}[X]} x\,p(x)$ "E" for expectation

$X \sim \text{Bern}(0.3)$ $\quad E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 1 \cdot 0.3 = 0.3$ (0·0.7 +)

$X \sim \text{Bern}(p)$
$E(X) = 0 \cdot p(0) + 1 \cdot p(1) = p(1) = \boxed{p}$

For general Bernoulli

3

completely independently ....

Why should $\bar{X} \to \mu$? Proof later...

Law of Large #s.

$X \sim \text{Binom}(8, 0.5) \Rightarrow$

We showed empirically (via dosmul data) that $E(X) \approx 4$

Let's prove it...

$\binom{8}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^4$ $\quad \binom{8}{8}\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^0$

$$E(X) = \sum_{x \in Supp(X)} x_i \, p(x_i) = 0 \cdot p(0) + 1 \cdot p(1) + 2\,p(2) + 3\,p(3) + 4\,p(4) + 5\,p(5) + 6\,p(6) + 7\,p(7) + 8\,p(8)$$

$$= 0 + 0.031 + 2 \cdot 0.109 + 3 \cdot 0.219 + 4 \cdot 0.273 + 5 \cdot 0.219 + 6 \cdot 0.109 + 7 \cdot 0.031 + 8 \cdot 0.004$$

$$= \boxed{4}$$

What about $X \sim \text{Binom}(n, p)$ $\quad E[X] = f(n, p)$? $= np$?

$$\sum_{x \in Supp(X)} x\, p(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = 0 + \sum_{x=1}^{n} \cdots = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} \cdots = \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} \cdots$$

$$= n \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\underbrace{(n-x)!}_{((n-1)-(x-1))}} \cdots = n \sum_{x=1}^{n} \binom{n-1}{x-1} p^x (1-p)^{n-x} = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{(n-1)-(x-1)}$$

let $m = n-1$
let $y = x-1$

$$= np \underbrace{\left(\sum_{y=0}^{m} \binom{m}{y} p^y (1-p)^{m-y} \right)}_{=1} = \boxed{np}$$

$x = 1 \ldots n$
$y = 0 \ldots n-1 = 0 \ldots m$

HARD:

$$X \sim \text{Hyper}(n, K, N) \quad E(X) = \sum_{x \in Supp(X)} x \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} = n\frac{K}{N} \text{ "}= np\text{" } ?$$

Wait... easier way...

$$X \sim \text{Unif}(1, 10, 100) \quad E(X) = \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 10 + \frac{1}{3} \cdot 100 = \frac{1}{3}(1 + 10 + 100) = \frac{111}{3}$$

$$X \sim \text{Unif}(A) \quad E(X) = \sum_{x \in A} x\, p(x) = \frac{1}{|A|} \sum_{x \in A} x$$

$X \sim \text{Geom}(p)$

$$\mu = E(X) = \sum_{x=1}^{\infty} x(1-p)^{x-1}p = \sum_{y=0}^{\infty}(y+1)(1-p)^{y}p = \sum_{y=0}^{\infty} y(1-p)^{y-1}p + \sum_{y=0}^{\infty}(1-p)^{y}p$$

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

$$= (1-p)\underbrace{\sum_{y=1}^{\infty} y(1-p)^{y-1}p}_{\mu} + 1 \Rightarrow \mu = (1-p)\mu + 1$$

$$\mu\left(1-(1-p)\right) = 1$$

$$\Rightarrow \boxed{\mu = \frac{1}{p}}$$

Is $E(X)$ the only useful summary of the r.v. X? No...

Let's review $p(x)$ and $F(x)$ again… and introduce "effective support"!

$X \sim Geom(0.2) = 0.8^{x-1} \cdot 0.2$

x	$p(x)$	$F(x)$
1	0.200	0.200
2	0.160	0.360
3	0.128	0.488
4	0.102	0.590
5	0.082	0.672
6	0.066	0.738
7	0.052	0.790
8	0.042	0.832
9	0.034	0.866
10	0.027	0.893
11	0.021	0.914
12	0.017	0.931
13	0.014	0.945
14	0.011	0.956
15	0.009	0.965
16	7	0.972
17	6	0.978
18	5	0.983
19	4	0.987
20	3	0.990
21	2	0.992
22	1	0.994
23	1	0.995
24	1	0.996
25	1	0.997
26	1	0.998
27	1	0.999
28	0.000	
⋮	Σ 0.999 ≈ 1	

$\bar{X} \to E[X]$ (LLN)

$E(X) := \sum_{x \in Supp(X)} x \, p(x)$

"prob weighted average"

$X \sim Bern(p)$

$E(X) = p$

$X \sim Binom(n, p)$

$E[X] = np$

$X \sim Hyper(n, K, N)$

$E(X) = n \frac{K}{N}$

(WAIT)

$X \sim Geom(p)$

$E(X)$?

$X \sim NegBin(r, p)$

$E(X)$?

$A \subset Supp[X]$

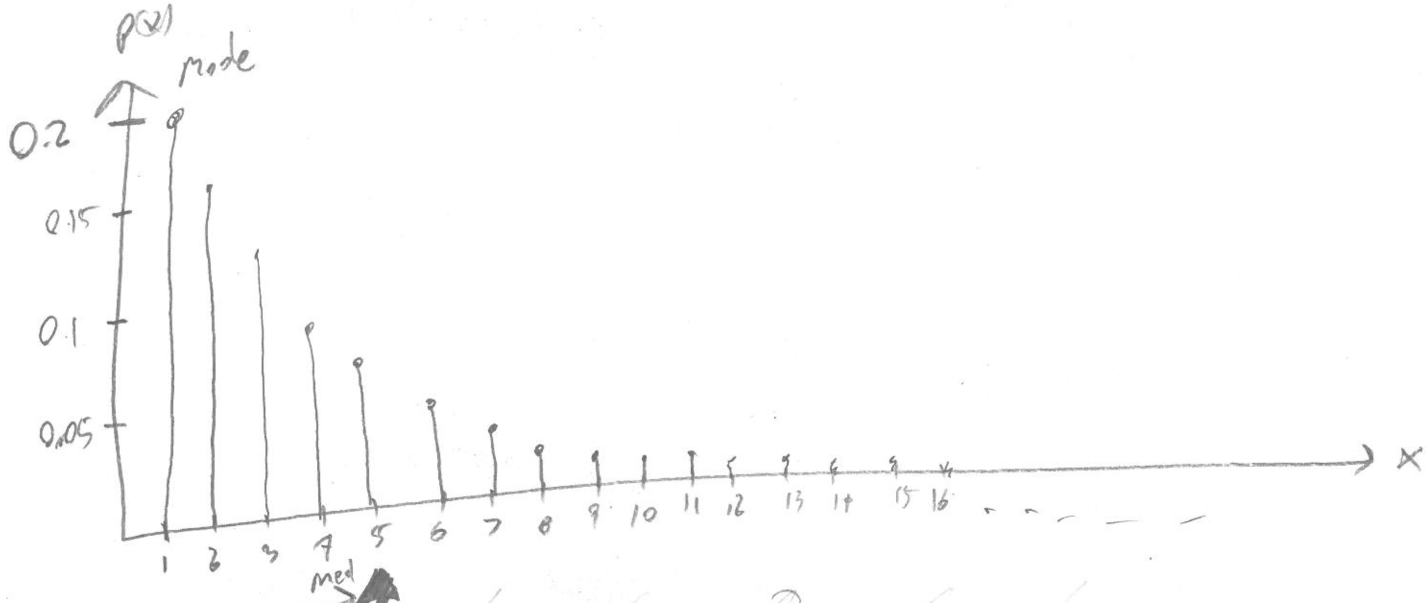
or smallest set s.t. $\sum_{x \in A} p(x) = 0.999$

Approximate / Effective Support

you choose

$\{x : x \in Supp(X) \text{ and } p(x) > 0.001\}$

why $\neq 1$?



Where's the balance pt? $E(X)$?

geometric series $= \frac{1}{p}$

$X \sim \text{Geom}(p)$ do the general case first

$$\mu = \sum_{x=1}^{\infty} x(1-p)^{x-1}p = \sum_{y=0}^{\infty}(y+1)(1-p)^y p = p\left(\sum_{y=0}^{\infty} y(1-p)^y + \sum_{y=0}^{\infty}(1-p)^y\right) = \sum_{y=0}^{\infty} y(1-p)^y p + 1$$

reminder: let $y = x-1 \Rightarrow x = y+1$
$x = 1 \ldots \infty \Rightarrow y = 0 \ldots \infty$

$$= (1-p)\underbrace{\sum_{y=1}^{\infty} y(1-p)^{y-1}p}_{\mu} + 1 \Rightarrow \mu = \mu - p\mu + 1 \Rightarrow \boxed{\mu = \frac{1}{p}}$$

$X \sim \text{Geom}(0.2) \Rightarrow \mu = \frac{1}{0.2} = 5$

$E[X]$ is a "functional" .. sums up r.v. with one #. Are there others that are useful?
$\text{Min}[X] = \min\{\text{Supp}[X]\}$, $\text{Max}[X] = \max\{\text{Supp}[X]\}$, $\text{Range}[X] = \text{Max}[X] - \text{Min}[X]$, $\text{Mode}[X] = \text{argmax}\{p(x)\}$

$$\text{Quantile}[X, p] := \underset{x \in \text{Supp}[X]}{\text{argmin}}\{F(x) \geq p\}$$ AKA "percentile" if p is a percent.

$\text{Quantile}[X, 0.5]$ is the "first" element in the support where 50% or more of the support is below it.

$\text{Median}[X] := \text{Quantile}[X, 0.5]$

~~$\text{IQR}[X] = Q[X, 0.75] - Q[$~~

Quantiles?

Ex $X \sim \text{Geom}(0.2)$

Quantile$[X, 0.25]$	2
Median(X)	4
Quantile$[Y, 0.75]$	7

$IQR(X) = Q[X, 0.75] - Q[X, 0.25] = 5$

Tertiles

Quantile$[X, \frac{1}{3}]$	2
Quantile$[X, \frac{2}{3}]$	5

Quintiles

$Q[X, 0.2]$
0.4
0.6
0.8

Deciles

$Q[X, 0.1]$ 1
0.2
⋮
$Q[X, 0.9]$ 11

Note: Median$(X) \neq E[X]$

Some Types of Distr.

Only in the case of "symmetric distr"

if Median$(X) < E[X]$ "skew right"

if $>$ "skew left"

If $|\text{mode}(X)| = 1$ "unimodal"

Custom r.v's

Roulette In America, bet on black. Payout 1:1

$$X \sim \begin{cases} \$1 & \text{w.p. } \frac{18}{38} \\ -\$1 & \text{w.p. } \frac{20}{38} \end{cases} \quad \mu = E[X] = \frac{-2}{38} = -\$0.053$$

Custom r.v. model we just build

If I play many times, my average winnings is $-\$0.053$

$\bar{X} \to \mu$ LLN

exp. value is only interpretable as a "long run" or large single payout

$T = X_1 + X_2 + \ldots + X_n$ $\quad \lim_{n \to \infty} T = -\infty$

If I play forever, I lose ∞!