1

Lecture 13 11/1/16 (exam 2 weeks from today)

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

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

| $x$ | $p(x)$ | $F(x)$ |
|---|---|---|
| 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 | |
| $\vdots$ | $\sum = 0.999 \approx 1$ | |

Review

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

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

"prob weighted average"

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

$E(X) = p$

$X \sim \text{Binom}(n, p)$

$E[X] = np$

$X \sim \text{Hyper}(n, K, N)$

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

(WAIT)

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

$E(X)$ ?

$X \sim \text{NegBin}(r, p)$

$E(X)$ ?

$A \subset Supp[X]$

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

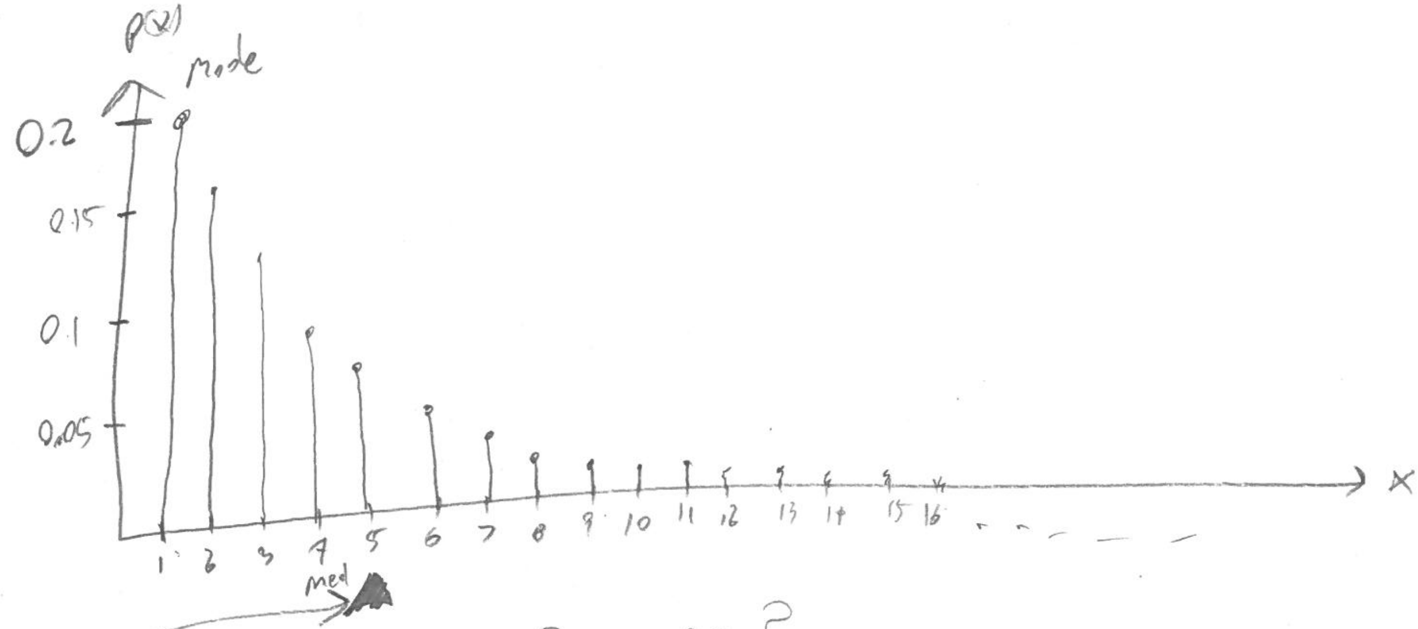
Approximate / Effective Support

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 $= \frac{1}{p}$

$X \sim Geom(p)$ 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 $y = x-1 \Rightarrow x = y+1$
$x = 1 \ldots \infty \Rightarrow y = 0 \ldots \infty$

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

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

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

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

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

$Median[X] := Quantile[X, 0.5]$

~~$IQR[X] = Q[X, 0.75] - Q1$~~

Quantiles?

Quantile [X, 0.25]

Median (X)

Quantile [Y, 0.75]

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

2

4

7

Tertiles

Quantile [X, 1/3] 2

Quantile [X, 2/3] 5

IQR[X] = Q[X, 0.75] - Q[X, 0.25] = 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 |mode[X]| = 1 "unimodal"

## Custom r.v's

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

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

Custom r.v. model we just built

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 sample property

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

If I play forever, I lose $\infty$!

Bet on lucky #7. Payout 35:1

$$X \sim \begin{cases} \$35 & \text{w.p. } \frac{1}{38} \\ -\$1 & \text{w.p. } \frac{37}{38} \end{cases}$$

Bet on Dozen 1-12. Payout 2:1

$$X \sim \begin{cases} \$2 & \text{w.p. } \frac{12}{38} \\ -\$1 & \text{w.p. } \frac{26}{38} \end{cases}$$

All bets the same expectation. Story...

Europe Bet on black. Same payout but...

$$X \sim \begin{cases} \$1 \\ -\$1 \end{cases} \qquad E[X] = -\$0.027 \text{ "much fairer"}$$

Def: "Fair Game" $E[X] = 0$

<u>Basic r.v. Transformations</u>  Uber example

$P(\text{traffic}) = 0.3$ If traffic → street, else Van Wyck

$$W \sim \begin{cases} 7\text{ min} & \text{w.p. } 0.7 \\ 12\text{ min} & \text{w.p. } 0.3 \end{cases} \quad \text{(not a Bernoulli)}$$

$$E[W] = 0.7 \cdot 7\text{ min} + 0.3 \cdot 12\text{ min} = 8.5\text{ min}$$ Interpret...

Over many trips, my average time in the taxi is $\approx$ 8.5 min. But any given trip...